

QY 353 PROARXHGHSGLUASPMETThrleu 361
 DB 664 CCCCAGCAGAGACATGACCTGCTA 690

RESULT 15
 BG715206
 LOCUS 602675668F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4798145 5',
 DEFINITION mRNA sequence.
 ACCESSION BG715206 702 bp mRNA linear EST_08-MAY-2001
 VERSION BG715206
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE 1 (bases 1 to 702)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@dbp-remail.nih.gov
 Tissue Procurement: Miklos Palokovics, M.D., Ph.D.
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shireki
 Toshitsuki and Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/ULNL at:
 http://image.llnl.gov
 Plate: LLM10685 row: h column: 18
 High quality sequence stop: 702.

FEATURES
 source location/qualifiers
 1..702
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4798145"
 /clone_lib="NIH MGC 96"
 /tissue_type="hypothalamus"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pBluescript (modified
 pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
); Oligo-dT primed using primer 5'-TTTTTTTATTATTN-3',
 size-selected for average insert size 2.3 kb and
 normalized to 5x. This is a primary library enriched
 for full-length clones and constructed using the
 Cap-trapper method (Carninci, in preparation). Library
 constructed by M. Brownstein (NIH/NHGRI, National
 Institutes of Health). Note: this is a NIH_MGC Library."

BASE COUNT 151 a 195 c 220 g 136 t
 ORIGIN

Alignment Scores:
 Pred. No.: 2.93e-105 Length: 702
 Score: 1129.00 Matches: 224
 Percent Similarity: 99.12% Conservative: 0
 Best Local Similarity: 99.12% Mismatches: 1
 Query Match: 43.76% Indels: 1
 DB: 12 Gaps: 0

US-09-830-144-4 (1-504) x BG715206 (1-702)

QY 1 MetAlAAGlnArgSerLeuGlnSerGlnGlnGlnProSerTrpThrAspAsp 20
 DB 26 ATGGCGGCGAGAGAGAGGAGTTCCTGCAGAGTGGAGCGAGCCAGACTGACAGATGAC 85

QY 21 LeuProLeuCyHisLeuSerGlyValGlySerAlaSerAsnArgSerTrpSerAlaAsp 40
 DB 86 CTGCCTCTGCGCACCTATCTGGGGTGGCTCAGCTCCAGCCGACGACTCTGCTGAT 145

QY 41 GlyLysGlyThrGlnSerHisProGluAspSerTrpLeuLysPheArgSerGluAsn 60

DB 146 GGCAGGGCACTGAGAGCCACCCGCCAGAGACAGCTGCCTCAAGTTCAGAGTCAAGAC 205
 QY 61 AasnCySphenLeuTrpGlyValPheAsnGlyTrpAspGlyAsnArgValThrAsnPheVal 80
 DB 206 AACTGCTCTCTGTATGGGGGTCTTCAACGCTATGATGAGCAACGAGTGCACCAACTTCGTG 265

QY 81 AlAGlnArgLeuSerAlaGlnLeuLeuGlyGlnLeuAsnAlaGlnHisAlaGlnAla 100
 DB 266 GCCCGAGCGGCTGTCCGAGAGCTCTCTGCGCACTGAAATGCCAGACGCCAGAGCC 325

QY 101 AspValArgArgValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGlnSer 120
 DB 326 GATGTGCGGCGGTGTGCTGTGCGAGGCTTCGATGTGTGAGAGAGGACTTCCTGAGATCC 385

QY 121 IleAspAspAlaLeuAlaGlnLysAlaSerLeuGlnSerGlnLeuProGlnGlyValPro 140
 DB 386 ATTTGACGACGCTTGCTGTGAGAAAGCAAGCTTCAGATTGCCCAATTCGACAGAGGAGTCCCT 445

QY 141 GlnHisGlnLeuProProGlnTrpGlnLysIleLeuGlnArgLeuLysThrLeuGluArg 160
 DB 446 CAGCACACAGCTGCTGCTCTCACTATCAAGATCTTGAGAGACTCAAGACGTTCAAGAGAG 505

QY 160 GGUlleserGlyGlyAlaMetAlaValAlaValAlaLeuLeuAsnAsnLysLeuTrpVal 180
 DB 506 GGAATTTCCGGAGGGGCGCATGGCGGTTGTGGCGTCTTCTCAACAAAGACTTACGT 565

QY 180 lAlaAsnValGlyThrAsnArgAlaLeuLeuCySlysSerThrValAspGlyLeuGlnVal 200
 DB 566 CCCCAATGCTGGTACAAACCGTGCACTTTATGCAATGAGACAGTGGATGGTTCAGAGT 625

QY 200 lThrGlnLeuAsnValAspHisThrThrGlnAsnGlnAspGlnLeuPheArgLeuSerGln 220
 DB 626 GACACAGCTGAACGTGGACCAACCAACAGAGAGATGAGCTCTTCCGCTTTCGCA 685

QY 220 lneGlyLeuAspAla 225
 DB 686 GCTGGGCTTGGATGCT 701

Search completed: December 10, 2002, 02:04:05
 Job time : 1587.89 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Inc.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 9, 2002, 23:00:44 (Search time 53.518 Seconds
(without alignments)
2886.095 Million cell updates/sec

Title: US-09-830-144-4
Perfect score: 2580
Sequence: 1 MAQRSLQSEQPSWTD.....AEFRLMSVDHGEOSVTAP 504

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODBL=frame+ p2n model -DEV=xld
-Q=/cgnt2.1/USPTO_spool/US09830144/runat_04122002_141354_2283/app_query.fasta.1.1422
-DB=Issued_Patents_NA -OPMT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdt
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09830144@CGN 1.1 44 @runat_04122002_141354_2283 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Issued_Patents_NA.*

1: /cgnt2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgnt2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgnt2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgnt2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgnt2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgnt2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2580	100.0	1560	2 US-08-752-891-1	Sequence 1, Appl1
2	2580	100.0	1560	2 US-09-144-178-1	Sequence 1, Appl1
3	2580	100.0	1560	3 US-09-406-854-1	Sequence 1, Appl1
4	2580	100.0	1560	4 US-09-529-279-1	Sequence 1, Appl1
5	2580	100.0	1560	4 US-09-529-279-10	Sequence 1, Appl1
6	2577	99.9	1560	4 US-09-529-279-42	Sequence 42, Appl1
7	2575	99.8	1560	2 US-08-752-891-5	Sequence 5, Appl1
8	2575	99.8	1560	2 US-09-144-178-5	Sequence 5, Appl1
9	2575	99.8	1560	3 US-09-406-854-5	Sequence 5, Appl1
10	178	6.9	1403	3 US-09-013-881-10	Sequence 10, Appl1
11	131	5.1	49377	1 US-08-764-233A-1	Sequence 1, Appl1
12	128.5	5.0	1824	2 US-08-822-701-1	Sequence 1, Appl1

13	128.5	5.0	1824	3 US-08-935-885-1	Sequence 1, Appl1
14	128.5	5.0	2081	3 US-08-935-885-21	Sequence 21, Appl1
15	126.5	4.9	28958	3 US-08-258-261B-6	Sequence 6, Appl1
16	126.5	4.9	28958	1 US-08-456-837-6	Sequence 6, Appl1
17	126.5	4.9	28958	1 US-08-457-342-6	Sequence 6, Appl1
18	126.5	4.9	28958	1 US-08-457-646A-6	Sequence 6, Appl1
19	126.5	4.9	28958	1 US-08-458-076A-6	Sequence 6, Appl1
20	126.5	4.9	28958	1 US-08-764-233A-4	Sequence 4, Appl1
21	126.5	4.9	28958	1 US-08-457-335A-6	Sequence 6, Appl1
22	126.5	4.9	28958	1 US-08-729-214-6	Sequence 6, Appl1
23	126.5	4.9	28958	3 US-08-028-934-6	Sequence 6, Appl1
24	125.5	4.9	1890	3 US-08-935-885-19	Sequence 19, Appl1
25	121	4.7	68750	3 US-09-335-409-1	Sequence 1, Appl1
26	121	4.7	68750	4 US-09-568-102-1	Sequence 1, Appl1
27	121	4.7	68750	4 US-09-567-969-1	Sequence 1, Appl1
28	121	4.7	68750	4 US-09-568-480-1	Sequence 1, Appl1
29	121	4.7	68750	4 US-09-568-486-1	Sequence 1, Appl1
30	121	4.7	68750	4 US-09-567-899-1	Sequence 1, Appl1
31	121	4.7	68750	4 US-09-567-899-1	Sequence 1, Appl1
32	117	4.5	3182	1 US-08-188-582-12	Sequence 12, Appl1
33	117	4.5	3182	1 US-08-646-715-12	Sequence 12, Appl1
34	117	4.5	71989	4 US-09-443-501A-2	Sequence 2, Appl1
35	116.5	4.5	4403765	4 US-09-103-840A-2	Sequence 2, Appl1
36	116.5	4.5	4411529	4 US-09-103-840A-1	Sequence 1, Appl1
37	116	4.5	1785	2 US-08-910-856-1	Sequence 1, Appl1
38	116	4.5	3187	2 US-08-910-856-9	Sequence 9, Appl1
39	116	4.5	3187	2 US-08-910-856-10	Sequence 10, Appl1
40	115.5	4.5	4660	4 US-09-166-350-26	Sequence 26, Appl1
41	115	4.5	4403765	4 US-09-103-840A-2	Sequence 2, Appl1
42	115	4.5	4411529	4 US-09-103-840A-1	Sequence 1, Appl1
43	113.5	4.4	44377	2 US-08-804-227C-7	Sequence 7, Appl1
44	113.5	4.4	44377	2 US-08-804-199-1	Sequence 1, Appl1
45	111.5	4.3	11601	2 US-08-222-617A-3	Sequence 3, Appl1

ALIGNMENTS

RESULT 1
US-08-752-891-1
: Sequence 1, Application US/08752891
: Patent No. 5837819
: GENERAL INFORMATION:
: APPLICANT: MATSUMOTO, Kunihiko
: APPLICANT: NISHIDA, Ritsuke
: TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR
: NUMBER OF SEQUENCES: 8
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Foley & Lardner
: STREET: 3000 K Street, N.W., Suite 500
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20007-5109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/752,891
: FILING DATE: 20-NOV-1996
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 8-300856
: FILING DATE: 28-OCT-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 8-126282
: FILING DATE: 24-APR-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: BENT, Stephen A.
: REGISTRATION NUMBER: 29,768
: REFERENCE/DOCKET NUMBER: 17961/111
: TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1560 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 30..1541
FEATURE: mat_peptide
NAME/KEY: mat_peptide
LOCATION: 30..1541
US-08-752-891-1

Alignment Scores:
Pred. No.: 4, 82e-249 Length: 1560
Score: 2580.00 Matches: 504
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-830-144-4 (1-504) x US-08-752-891-1 (1-1560)

QY 1 MetAlaAlaGlnArgSerLeuLeuGlnSerGluGlnProSerThrThrAsp 20
DB 30 ATGGCGGCGCAGAGAGAGCTTGTGCAGAGTGAGCAGCAGCCCAAGCTGGACAGATGAC 89
QY 21 LeuProLeuCysHisLeuSerGlyValGlySerAlaSerAsnArgSerThrAlaAsp 40
DB 90 CTGCCCTCTCTCCCACTCTCTGGGGTTGGCTCAGCCTCCCAACCGCAGCTACTCTGCTGAT 149
QY 41 GlyLysGlyThrGluSerHisProGluAspSerTrpLeuLysPheArgSerGluAsn 60
DB 150 GCGAAGGCGACTGAGAGCCACCCCGCAGAGGACAGCTGGCTCAAGTTCCAGGAGTGAGAAC 209
QY 61 AsnCysPheLeuValGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsnPheVal 80
DB 210 AACTGCTTCTCTGATGGGCTTCAACGGCTATGATGGCAACCGAGTGACCACTTCGTG 269
QY 81 AlaGlnArgLeuSerAlaGluLeuLeuLeuGlyGlnLeuAsnAlaGluHisAlaGluAla 100
DB 270 GCCCAGCGGCTGTCCGACAGAGCTCTGCTGGCCAGCTGAATGCCGACGACGCCGCGGCC 329
QY 101 AspValArgArgValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGluSer 120
DB 330 GATGTGCGGCGTGTCTGTCTGAGGCTTCGATGTGGTGAGAGAGCTTCTTGAGTCC 389
QY 121 IleAspAlaLeuAlaGluLysAlaSerLeuGlnSerGlnLeuProGluGlyValPro 140
DB 390 ATTGACGACGCTTGGCTGGAAGGCAAGCCTCCAGTCCCAATTCGACAGAGGAGTCCCT 449
QY 141 GlnHisGlnLeuProGlnTyrGlnLysIleLeuGluArgLeuLysThrLeuGluArg 160
DB 450 CAGCACCACTGCTCTCTAGTATCAGAGATCTTTGAGAGACTCAAGACCTTAGAGAGG 509
QY 161 GluLysGlyGlyAlaMetAlaValAlaValLeuLeuAsnAsnLysLeuTyrVal 180
DB 510 GAAATTTGGGAGGGGCGCATGGCGGTGTGGCGGTCTTCTCAACCAACCAAGCTTACGTC 569
QY 181 AlaAsnValGlyThrAsnArgAlaLeuLeuCysLysSerThrValAspGlyLeuGlnVal 200
DB 570 GCCAATGTCGTACAAACCGTGCACCTTTATGCAATCAGACAGTGTGATGGTTCAGGTG 629
QY 201 ThrGlnLeuAsnValAspHisThrThrGluAsnGluAspGluLeuPheArgLeuSerGln 220
DB 630 ACACAGCTGAACGTGGACCAACACACAGAGACGAGATGAGCTTCTCCGCTTTCGACG 689
QY 221 LeuGlyLeuAspAlaGlyLysIleLysGlnValGlyIleLysCysGlyGlnGluSerThr 240

DB 690 CTGGCTTGGATGCTCGAAAGATCAAGCAGGTGGGATCATCTGTGGCGAGGAGCACC 749
QY 241 ArgArgIleGlyAspTyrLysValLysTyrGlyTyrThrAspIleAspLeuLeuSerAla 260
DB 750 CGCGGATCGGGATTACAAAGTTAAATATGGCTACAGGACATTGACCTTCTCAGCGCT 809
QY 261 AlalysSerLysProIleAlaGluProGluIleHisGlyAlaGlnProLeuAspGly 280
DB 810 GCCAAGTCCAAACCAATCATCGACAGCCAGAAATCCATGGGGCACAGCCGCTGGATGG 869
QY 281 ValThrGlyPheLeuValLeuMetSerGluGlyLeuTyrLysAlaLeuGluAlaHis 300
DB 870 GTGACGGGCTTCTGGTGTCTGATGTCGAGGGGTGTACAGGGCCCTAGAGGAGCCCAT 929
QY 301 GlyProGlyGlnAlaAsnGlnGluAlaMetIleAspThrGluPheAlaLysGln 320
DB 930 GGGCTGGCGAGGCCAACCCAGGAGATTGCTGCGATGTTGACACTGAGTTTGCACAGCAG 989
QY 321 ThrSerLeuAspAlaValAlaGlnAlaValValAspArgValLysArgIleHisSerAsp 340
DB 990 ACCTCTCTGGACGAGTGGCCCGAGCCCTCTGGACCGGGTGAAGCGCATCCACAGCGAC 1049
QY 341 ThrPheAlaSerGlyGlyGluArgAlaArgPheCysProArgHisGluAspMetThrLeu 360
DB 1050 ACCTTCGCCAGTGTGGGAGCGTCCAGGTTCTGCCCGCGCACGAGGACATGACCTG 1109
QY 361 LeuValArgAsnPheGlyTyrProLeuGlyGluMetSerGlnProThrProSerProAla 380
DB 1110 CTAGTGAGGAACCTTGGCTACTCCCGCTGGCGAATGAGCCAGCCACACCGAGCCGACCC 1169
QY 381 ProAlaAlaGlyGlyArgValTyrProValSerValProTyrSerSerAlaGlnSerThr 400
DB 1170 CCAAGTGCAGGAGGAGTGTACCTGTGTGTGCTACTCCAGCGCCGAGAGCACC 1229
QY 401 SerLysThrSerValThrLeuSerLeuValMetProSerGlnGlyGlnMetValAsnGly 420
DB 1230 AGCAAGACCAAGCTGACCTCTCTCTGTCATGCCCTCCAGGGGCCAGATGGTCAACGG 1289
QY 421 AlaHisSerAlaSerThrLeuAspGluAlaThrProThrLeuThrAsnGlnSerProThr 440
DB 1290 GCTCACAGTCTTCCACCTTGAGGAGCCACCCCTCCACCAACCAAGCCGACCC 1349
QY 441 LeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSerSerAspGly 460
DB 1350 TTAACCTGTCAGTCCACCAACACGACACGACAGCAGCAGCTCCAGCTCTGACGGAGCC 1409
QY 461 LeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGlyArgValGluPro 480
DB 1410 CTCTTCCGCTCCCGGCCGCCACTCGCTCCGCTGGCGAGGAGCGTCTGTGTGAGCCC 1469
QY 481 TyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGlyGlnSerVal 500
DB 1470 TATGTGACTTTGCTGAGTTTACCGCTCTGGAGCGTGGACCATGGCGAGCAGAGCGTG 1529
QY 501 ValThrAlaPro 504
DB 1530 GTGACAGCACC 1541

RESULT 2

US-09-144-178-1
; Sequence 1, Application US/09144178
; Patent No. 5989862
; GENERAL INFORMATION:
; APPLICANT: MATSUMOTO, Kunihiro
; APPLICANT: NISHIDA, Eisuke
; TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESSES:
; ADDRESSES: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA

ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/144,178
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/752,891
FILING DATE: 20-NOV-1996
APPLICATION NUMBER: JP 8-300856
FILING DATE: 28-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 8-126282
FILING DATE: 24-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 17981/111
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ. ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1560 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 30..1541
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 30..1541
US-09-144-178-1
Alignment Scores:
Pred. No.: 4.82e-249 Length: 1560
Score: 2580.00 Matches: 504
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
US-09-830-144-4 (1-504) x US-09-144-178-1 (1-1560)

Qy 121 Ileaspaalaleualaglulysalaserleuenglnserglnleuproglnlyvalpro 140
Db 390 ATTGACGAGCGCTTGGCTGAGAGGCAAGCTCCAGTGCATTTGCCAGAGAGCTCCCT 449
Qy 141 Glnhsglnleuproproglnltyrghlnlylleyengluargleulythrleugluarg 160
Db 450 CAGCACCGAGCTGCTCTCCAGATCAGAAAGATCCTTGAGAACTCAAGACCTTGAAGAGG 509
Qy 161 GlnleserghlygylalameclavalavalalleuleuAasnleuTyrrval 180
Db 510 GAAATTTGGAGGGGGCCATGAGCGCTTGGCGGCTCTTCAACACAACTCTACGTC 569
Qy 181 AlaenvalgllythrAasnarglaleuleuCyshysserthrvalaspglyleuglnval 200
Db 570 GCCATGTGCGTACAAACCGTGCACCTTTATGCAAAATGACAGTGGATGGGTTCAGGTG 629
Qy 201 ThrghlnleuAenvalasphisthrghlnuansglnuaspghuleuPheargleuSerghn 220
Db 630 ACACAGCTGAACGTGACACACACACAGAGAAACGAGATGAGCTCTTCGTTCCGAG 689
Qy 221 leuaglyleuaspalagllyslleyghlnvalglylleCyshlyghlnghlserthr 240
Db 690 CTGGGCTTGATGCTGGAAGAAGTCAAGAGGTGGGATCATCTGTGGCAGAGAGCACCC 749
Qy 241 ArgarglleglyAspyrlysevalyslyrghlyrthrAspgleuSerleuSerAla 260
Db 750 CGCGGATGGGGATTACAGGTTAAATATGCTACAGGACATTGACCTTCAGCGCT 809
Qy 261 AlalyserlyserproillealagluproglnlehslygylalaglnProleuaspghly 280
Db 810 GCCAAGTCCAAACCAATCATCCAGAGGCGAGAAATCCATGGGGCACACCCCTGGATGG 869
Qy 281 ValThrghlyPheleuValleuwerSerghlyleuTyrrlysalaleuGlnAlaAlahis 300
Db 870 GTGACGGGCTTCTGTGCTGATGTGGAGGGTGTTCAGAGCCCTTGAGAGGACCCAT 929
Qy 301 GlyproghlyghlnalAasnghlnlullealameclleasprthrghlPhealalyghln 320
Db 930 GGGCTGGGCAAGCCACAGAGAGATGCTGCGATGATTCACCTGATTTGCCAAGCAG 989
Qy 321 ThrSerleuaspAlaValalaglnalavalalaspargvalysargillehsersp 340
Db 990 ACCTCCCTGAGCGAGTGGCCAGGCCGCTGCGAGCCGGGTGMAACCGCATCCACAGCCAC 1049
Qy 341 ThrPheAlaserghlyghlnuarglaargPheCysproarghlysluaspmetthrleu 360
Db 1050 ACCTTGCCAGTGGTGGAGCGTGCAGGTTCTGCCCCGACAGAGACATGACCTTG 1109
Qy 361 leuValIargAsnPhesgllyrProleuenglylumetSerghnProThrProserProAla 380
Db 1110 CTAGTAGAAGACTTGGCTACCCGCTGGCGAATGAGCCAGCCACAGCCAGCC 1169
Qy 381 ProAlaAlaglyghlyargvallyrProvalserValProTySerSerAlaglnSerThr 400
Db 1170 CCAGCTGACAGAGAGAGAGATGATACCTGTGTGTGCTACTCCAGGCGCCAGAGCAC 1229
Qy 401 SerlyThrSerValthrleuSerleuValmetProserghlnlyGlnmetValasnghly 420
Db 1230 AGCAAGACCAAGCTGACCTCTCCCTGTGATGCTCCACAGGCGCAGATGATCAAGGG 1289
Qy 421 AlaHsserAlaserthrleuaspghlnalathrProthrleuThrAsnghlnserProThr 440
Db 1290 GTCACAGTGTCTTCCACCTTGACGAGACACCCCACTCCACCAACCAAGGCCGAC 1349
Qy 441 leuThrleuGlnserThrAsnthrhisthrghlnserSerSerSerSeraspghlygly 460
Db 1350 TTAACCTGAGTCCACCAACACGACAGCAGAGAGCAGCTCCAGCTTACAGAGGC 1409
Qy 461 leuPheargSerthrproAlahisserleuProproglnlyGluAspghlyargvalghlpro 480
Db 1410 CTCTTCGCTCCCGGCCCGCCACTGCTCCGCTTGGCAGAGAGCGGTGTTAGGCC 1469

Qy 481 TyrValAspPheAlaGluPheTyrArgLeuTyrSerValAspHisGlyGluGlnSerVal 500
Db 1470 TATGTGACCTTTGCTGAGTTTACCGCTCTGGAGCGTGGACCATGGCGAGAGCGTG 1529
Qy 501 ValThrAlaPro 504
Db 1530 GTGACAGCACCG 1541

RESULT 3
US-09-406-854-1
; Sequence 1, Application US/09406854
; Patent No. 6140042
; GENERAL INFORMATION:
; APPLICANT: MATSUMOTO, Kunihiro
; APPLICANT: NISHIDA, Eisuke
; TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/406,854
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/752,891
; FILING DATE: 20-NOV-1996
; APPLICATION NUMBER: JP 8-300856
; FILING DATE: 28-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-126282
; FILING DATE: 24-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 17981/111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1560 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 30..1541
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 30..1541
US-09-406-854-1

Alignment Scores:
Pred. No.: 4,828-249 Length: 1560
Score: 2580.00 Matches: 504
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-830-144-4 (1-504) x US-09-406-854-1 (1-1560)

Qy 1 MetAlaAlaGlnArgArgSerLeuLeuGlnSerGluGlnGlnProSerTyrThrAspAsp 20
Db 30 ATGGCGGCGCAGAGAGAGAGCTTCTGCAGAGTGGAGCAGCAAGCTGGACAGATCAC 89
Qy 21 LeuProLeuCysHisLeuSerGlyValGlySerAlaSerAsnArgSerTyrSerAlaAsp 40
Db 90 CTGCGCTCTCTGCCACCTCTCTGGGGTGGCTCAGCCTCCAAACCGCAGCTACTCTGTGAT 149
Qy 41 GlyLysGlyThrGluSerHisProGluAspSerTrpLeuLysPheArgSerGluAsn 60
Db 150 GGCAGGGCAGCTGAGAGCCACCCGCCAGAGCAGCTGGCTCAAGTTTCAGGAGTGAAC 209
Qy 61 AsnCysPheLeuTyrGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsnPheVal 80
Db 210 AACTGCTTCTCTGCCACCTCTCTGGGGTGGCTCAGCCTCCAAACCGCAGCTACTCTGTGAT 269
Qy 81 AlaGlnArgLeuSerAlaGluLeuLeuGlyGlnLeuAsnAlaGluHisAlaGluAla 100
Db 270 GCCCAGCGGCTGTCCGAGAGCTCTCTGCTGGGCCAGCTGAATGCCAGCAGCGCCGAGGCC 329
Qy 101 AspValArgArgValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGluSer 120
Db 330 GATGTGCGCGTGTGCTGTCAGGCGCTTCGATGTGGTGGAGAGAGCTTCTCTGGAGTCC 389
Qy 121 IleAspAspAlaLeuAlaGluLysAlaSerLeuGlnSerGlnLeuProGluGlyValPro 140
Db 390 ATTGACGACGCTTGGCTGAGAGCGCAAGCTTCCAGTCGCAATTGCCAGAGGGATCCCT 449
Qy 141 GlnHisGlnLeuProGlnTyrGlnLysIleLeuGluArgLeuLysThrLeuGluArg 160
Db 450 CAGCACCGAGCTGCTCTCTCAGTATCAGAGATCTCTGAGAGACTCAAGAGCTTAGAGAGG 509
Qy 161 GluIleSerGlyGlyAlaMetAlaValAlaValLeuLeuAsnLysLeuTyrVal 180
Db 510 GAAATTTCCGGAGGGGCGCATGGCGTGTGGCGTCTTCTCAACAACAAGCTCTACGTC 569
Qy 181 AlaAsnValGlyThrAsnArgAlaLeuLeuCysLysSerThrValAspGlyLeuGlnVal 200
Db 570 GCCAATGTGGGTACAAACCGTGCACTTTTATGCAAAATCGACAGTGGATGGGTGGAGGTG 629
Qy 201 ThrGlnLeuAsnValAspHisThrThrGluAsnGluAspGluLeuPheArgLeuSerGln 220
Db 630 ACACAGCTGAACGTGGACACACACAGAGAACGAGATGAGCTCTTCCGCTCTTCGAG 689
Qy 221 LeuGlyLeuAspAlaGlyLysIleLysGlnValGlyIleLeuLysGlyGlnGluSerThr 240
Db 690 CTGGGCTTGGATGCTGGAAGATCAAGCAGTGGGGATCATCTGTGGCAGGAGAGCACC 749
Qy 241 ArgArgIleGlyAspTyrLysValLysValLysTyrGlyTyrThrAspIleAspLeuSerAla 260
Db 750 CGCGGATCGGGATTACAAGGTAAATATGGCTACACGGACATTGACCTTCTCAGCGCT 809
Qy 261 AlaLysSerLysProIleIleAlaGluProGluIleHisGlyAlaGlnProLeuAspGly 280
Db 810 GCCAAGTCCAAACCAATCATCGCAGAGCCAGAAATCCATGGGGCAGACGCGCTGGATGGG 869
Qy 281 ValThrGlyPheLeuValLeuMetSerGluGlyLeuTyrLysAlaLeuGluAlaHis 300
Db 870 GTGACGGGCTTCTTGGTGTGATGTGGAGGGTGTGTACAAGGCCCTTAGAGGCGAGCCCAT 929
Qy 301 GlyProGlyGlnAlaAsnGlnGluIleAlaAlaMetIleAspThrGluPheAlaLysGln 320
Db 930 GGGCTGGGCGAGGCCAACCCAGGAGATTGCTGCGATGATTGACACTGAGTTTGCACAGCAG 989
Qy 321 ThrSerLeuAspAlaValAlaGlnAlaValValAspArgValLysArgIleHisSerAsp 340
Db 990 ACCTCCCTGGACGAGTGGGCCCGCCAGGCGCTGTGGACCGGGTGAAGCGCATCCACACGGAC 1049
Qy 341 ThrPheAlaSerGlyGlyGluArgAlaArgPheCysProArgHisGluAspMetThrLeu 360
Db 1050 ACCTTCGCCAGTGGTGGGAGCGTGCAGGTTCTGCCCCCGGACGAGACATGACCCCTG 1109
Qy 361 LeuValArgAsnPheGlyTyrProLeuGlyGluMetSerGlnProThrProSerProAla 380

Db	1110	CTAATGAGGAACCTTTGGCTAACCCGCTGGCGGAAATGACCAAGCCACCAAGCCACGACC	1169
QY	381	ProAlaIaIaGlyIaArgVal ¹ TyrProValSerVal ¹ ProTyrSerSerAl ¹ GlnSerThr	400
Db	1170	CCAGCTCAGAGAGACGACGAGTACCTCTGTCTGTGCCATCTCCAGAGGCCACAGACACC	1229
QY	401	SerIysThrSerVal ¹ ThrIeuSerIeuVal ¹ MetProSerGln ¹ GlnMetVal ¹ AsnGly	420
Db	1230	AGCAAGACCAAGCGTGAACCTCTCCCTTGTTCATGCCCTCCAGAGGCCAGATGGTCAACGGG	1289
QY	421	AlaHisSerAlaSerThrIeuAspGlu ¹ AlaThrProThr ¹ IeuThr ¹ AsnGlnSerProThr	440
Db	1290	GCTCACAGTGTCTTCCACCTCGAGACGAAGCACCCACCTCACCAACCAAGCCGACACC	1349
QY	441	IeuThrIeuGlnSerThrAsnThr ¹ AsnGlnSerSerSerSerSerSerSerSerSerSer	460
Db	1350	TTAACCCCTGCAGTCCACCAACAGCAACGACGAGCAGAGCTTCAGCTTCAGCGAGGCC	1409
QY	461	IeuPheArgSerArgProAlaHisSerIeuPro ¹ GluIuAspGly ¹ ArgVal ¹ GluPro	480
Db	1410	CTCTTCGCGCTCCGCGCCCGCCCACTGGCTCCGCGCTCGGAGAGACGGTGTGTTAAGGCC	1469
QY	481	TyrVal ¹ AspPheAlaGluPhe ¹ TyrArgIeu ¹ TrpSer ¹ Val ¹ AspHisGly ¹ GlnGlnSerVal	500
Db	1470	TATGTGGACTTGTGCTGAGTTTAAACCGGCTCTGGAGCGTGGACCAATGACGACAGAGCTGTG	1529
QY	501	Val ¹ ThrAlaPro ¹ 504	
Db	1530	GTGACAGCACCG 1541	

```

RESULT 4
US-09-529-279-1
; Sequence 1, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF- $\beta$  INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529, 279
; CURRENT FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1560
; TYPE: DNA
; ORGANISM: Homo-sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (30)..(1541)
US-09-529-279-1

Alignment Scores:
Pred. No.: 4,82e-249 Length: 1560
Score: 2580.00 Matches: 504
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-830-144-4 (1-504) x US-09-529-279-1 (1-1560)

QY 1 MetAlAlAGlnArgSerLeuGlnSerGlnGlnProSerTyrThrAsp 20
Db 30 ATGGCGGGCGAGGAGGAGCTTCTGCTCGAGATGACACACACCAAGCTGACACATGAC 89
QY 21 LeuProLeuGlnSerLeuSerGlyValAlGlySerAlaSerAsnArgSerTyrSerAlaAsp 40

```

Db 90 CTGCCTCTCTGCCACTCTCTGGGGTTGGCTCAGCTCCAAAGCCAGACTACTCTGTAAT 149

QY 41 GLIYVSGIYThrgIuSerH3ProProGIuASpSerTriPleuIuSPhaeIuSergIuAen 60

Db 150 GGCAAGGCACTGAGAGCCAAAGCCGCAAGAGACAGCTGAGCTCAAGTTCAAGATGAGAAC 209

QY 61 AencSPhaeIuSryGIyValPheAenGIYTYrAspGIyAsnArGIyAlThraSPhaeVal 80

Db 210 AACTGCTCTCGIATGGGGTCTTCAACGGCTATGATGCAACGAGTGAACCACTTCGTA 269

QY 81 AIAgIuArGIuSerAlaGIuIuLeuIuLeuGIuIuLeuAenAlaGIuIuSAlaGIuAla 100

Db 270 GCCCAGCGCGTCTGCCAGAGCTCTCTGGCGCAGCTGAATGCCAGACGCCGAGGCC 329

QY 101 AspValArGIrGIyValIleuLeuGlnAlaPheAspValValGIuArGSePheLeuGIuSer 120

Db 330 GATGCGCGGCGTGTCTGTCTGACAGCCTTCGATGTGGTGAAGAGAGACTTCCTGGAGTCC 389

QY 121 ILeaSpAlaIuLeuAlaGIuIuValISeIuLeuIuSergIuLeuProGIuGIyValPro 140

Db 390 ATTGACGACGCTTGGCTGTGAGAAAGCAAGCCTCCAGTGCAGATTGCCAGAGGAGTCCCT 449

QY 141 GlnHISgInIuLeuProProGIuIuTYrGIuIuSAlIeLeuGIuArGIuLeuIuSPhaeIuAriG 160

Db 450 CAGCACCGAGCTGCTCCCTCACTATCAAGAGATCTTTGAGAGACTCAAGAGCTTAGAGAGG 509

QY 161 GIuIuSergIyGIyAlaMeArIaValIaValaValIleuIuSnaIuSleuIuTYrVal 180

Db 510 GAAATTTGGAGAGGGGCATAGCCGTTGTGGCGGCTCTTCTCAACAAACAAAGCTTACGTC 569

Qy	181	Ala	Asn	Val	Gly	Thr	Asn	Ala	Ala	Leu	Leu	Cys	Val	Ser	Thr	Val	Asp	Gly	Leu	Ala	200				
Db	570	GCC	ATG	CGG	TAT	CAAA	CCG	GC	ACT	TTT	AT	GCA	AT	TCA	GAC	AGT	GGA	TGG	ATT	G	629				
Qy	201	Thr	Gln	Leu	Asn	Val	Asp	His	Ser	Thr	Thr	Gln	Val	Asn	Gln	Val	Asp	Gln	Leu	Phe	Arg	Leu	Ser	Gln	220
Db	630	ACA	CAG	CGT	GAA	CG	GGA	CC	AC	CAC	CA	GGA	AA	GAG	AGA	TAG	ACT	CTT	CCG	TTT	TTC	G	C	689	
Qy	221	Leu	Gly	Leu	Asp	Ala	Val	Gly	Val	Ser	Gln	Val	Gly	Leu	Leu	Cys	Gly	Gln	Gln	Leu	Ser	Thr			240
Db	690	CTG	GCT	CTT	GAT	TGC	TGA	AAG	ATC	AA	G	AC	GAG	TAC	GAG	TGG	TGG	GAT	CAT	CTG	TGG	G	C	749	
Qy	241	Arg	Arg	Gln	Leu	Gly	Asp	Tyr	Leu	Val	Ser	Tyr	Gly	Tyr	Thr	Asp	Gln	Leu	Ser	Ala					260
Db	750	CGG	CGG	ATC	GCG	GAT	TCA	AG	GTT	AAA	TAT	GCG	TAC	AC	GGA	CA	TTC	ATC	CTT	C	A	G	C	809	
Qy	261	Ala	Val	Ser	Leu	Ser	Pro	Ile	Ile	Ala	Gln	Pro	Gln	Ile	His	Gly	Val	Ala	Gln	Pro	Leu	Asp	Gly		280
Db	810	GCC	AA	GCT	CAAA	CC	AA	TCA	TAT	CG	TAC	GAG	CC	CA	AA	TCA	TAT	CGG	GC	CA	G	C	G	C	869
Qy	281	Val	Thr	Gly	Phe	Leu	Val	Leu	Met	Ser	Gln	Gly	Leu	Tyr	Leu	Val	Ala	Gln	Ile	Ala	His				300
Db	870	GTA	G	CGG	CTT	CTT	G	TG	TG	TG	ATG	TG	CGA	AG	G	GGT	TTC	TA	CA	AG	CC	T	C	T	929
Qy	301	Gly	Pro	Gln	Gly	Ala	Asn	Gln	Gln	Ile	Ala	Ala	Met	Ile	Asp	Thr	Gln	Phe	Ala	Ser	Gln				320
Db	930	GGG	CTT	G	GC	AGG	CC	AA	CC	AG	AG	AT	TG	CT	G	CA	TAT	TAC	CA	CT	AG	T	T	T	989
Qy	321	Thr	Ser	Leu	Asp	Ala	Val	Ala	Gln	Ala	Val	Ala	Asp	Arg	Val	Ala	Ser	Arg	Ile	His	Ser	Asp			340
Db	990	ACC	TCC	CTG	AG	CG	CAG	TGG	CC	CA	GCG	CG	T	GT	GGA	CC	G	GGT	GT	AA	G	C	T	C	1045
Qy	341	Thr	Phe	Ala	Ser	Gly	Gln	Arg	Ala	Arg	Phe	Cys	Pro	Arg	His	Gln	Val	Ser	Met	Thr	Leu				360
Db	1050	ACC	TTG	CGC	AGT	GAT	G	GG	AG	CG	TCC	CA	G	GT	CTT	CC	CC	CG	CA	CA	G	A	C	A	1105
Qy	361	Leu	Val	Arg	Asn	Phe	Gly	Tyr	Pro	Leu	Gly	Val	Met	Ser	Gln	Pro	Thr	Pro	Ser	Pro	Ala				380
Db	1110	CTA	GTG	AG	AG	AA	CTT	TGG	CTA	CC	CT	G	GG	AG	AA	TG	CA	G	CC	CA	C	A	C	C	1165
Qy	381	Pro	Ala	Ile	Arg	Gly	Arg	Val	Tyr	Pro	Val	Ser	Val	Pro	Tyr	Ser	Ser	Ile	Arg	Ins	Thr				400

Db 1170 CCAGCTGAGGAGCAGAGTGTACCTGTGTCTGTGCCATCTCCAGCGCCAGAGCACC 1229
Qy 401 SerLysThrSerValThrLeuSerLeuValMetProSerGlnGlnMetValAsnGly 420
Db 1230 AGCAAGACCAAGCGTGACCTCTCCCTTGTCTCATGCCCTCCAGGGCCAGATGGTCAACGGG 1289
Qy 421 AlaHisSerAlaSerThrLeuAspGluAlaThrProThrLeuThrAsnGlnSerProThr 440
Db 1290 GCTCACAGTGTCTCCACCTCGGAGCAAGCCACCCCACTCCACCAACCAAGCCCGACC 1349
Qy 441 LeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSerSerSerSerSer 460
Db 1350 TTAACCTGTGAGTCCACCAACGACACGACGAGAGAGAGCTCCAGCTCTGACGGAGGC 1409
Qy 461 LeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGlyArgValGluPro 480
Db 1410 CTCTTCGGCTCCCGGCCCGCCACTCGCTCCCGCTGGCGAGGACGGTCTGTGTGAGCCC 1469
Qy 481 TyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGlyGluGlnSerVal 500
Db 1470 TATGTGGACTTTTGTGAGTTTATCCGCCCTCTGGAGCGTGACCATGGCGAGCAGACGGT 1529
Qy 501 ValThrAlaPro 504
Db 1530 GTGACAGCACCG 1541

RESULT 5

US-09-529-279-10
; Sequence 10, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529, 279
; CURRENT FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JF98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 1569
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)..(1557)
US-09-529-279-10

Alignment Scores:
Pred. No.: 4, 87e-249 Length: 1569
Score: 2580.00 Matches: 504
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-830-144-4 (1-504) x US-09-529-279-10 (1-1569)

Qy 1 MetAlaAlaGlnArgArgSerLeuGlnSerGluGlnGlnProSerTrpThrAsp 20
Db 7 ATGGCGCGCAGAGGAGGAGCTGTGTGACAGTGGAGCAGCAAGCTGGACAGATGAC 66
Qy 21 LeuProLeuCyHisLeuSerGlyValGlySerAlaSerAsnArgSerTyrSerAlaAsp 40
Db 67 CTGGCTCTCTGCACCTCTCTGGGTTGGCTCAGCTCCACCGCAGCTACTCTCTCTGAT 126
Qy 41 GlyLysGlyThrGluSerHisProGluAspSerTrpLeuLysPheArgSerGluAsn 60

Db 127 GCAAGGGCACTGAGAGCCACCCGCCAGAGCAGAGCTGGTCAAGTTTCAGGAGTGAGAAC 186
Qy 61 AsnCysPheLeuTyrGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsnPheVal 80
Db 187 AACTGCTTCCTGTATGGGCTCTTCAACGGCTATGATGGCAACCGAGTGACCACTTCGTG 246
Qy 81 AlaGlnArgLeuSerAlaGluLeuLeuGlyGlnLeuAsnAlaGluHisAlaGluAla 100
Db 247 GCCACCGCGCTGTCCGACAGAGCTCTCTGGGCCAGCTGAATGCCGAGCAGCCGAGGCC 306
Qy 101 AspValArgArgValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGluSer 120
Db 307 GATGTGCCGCTGTGCTGTGTCAGGCTTCGATGTGTGGAGAGGAGCTCTCTGGAGTCC 366
Qy 121 IleAspAspAlaLeuAlaGluLysAlaSerLeuGlnSerGlnLeuProGluGlyValPro 140
Db 367 ATTGACGAGCGCTTGGCTGAGAGGCAAGCCTCCAGTCGCAATTTGCCAGAGGAGTCCCT 426
Qy 141 GlnHisGlnLeuProProGlnTyrGlnLysIleLeuGluArgLeuLysThrLeuGluArg 160
Db 427 CAGCACAGCTGCCCTCCCTCAGTATCAGAAGATCCTTCAGAGACTCAAGACGTTAGAGAGG 486
Qy 161 GluIleSerGlyGlyAlaMetAlaValAlaValLeuLeuAsnAsnLysLeuTyrVal 180
Db 487 GAAATTTCCGGAGGGGCCATGGCCGTTGTGGCGGTCCTTCTCAACAACAAGCTCTACGTC 546
Qy 181 AlaAsnValGlyThrAsnArgAlaLeuLeuCysLysSerThrValAspGlyLeuGlnVal 200
Db 547 GCCAATGTCCGTACAAACCGCTGACATTTTATGCNAATCGACAGTGGATGGGTTCAGGTG 606
Qy 201 ThrGlnLeuAsnValAspHisThrThrGluAsnGluAspGluLeuPheArgLeuSerGln 220
Db 607 ACACAGCTGAACGTGGACCAACACACAGAGAACGAGGATGAGCTCTTCCGCTCTTCGACG 666
Qy 221 LeuGlyLeuAspAlaGlyLysIleLysGlnValGlyIleLysCysGlyGlnGluSerThr 240
Db 667 CTGGCTTGGATCTGGAAAGATCAACGAGGTGGGGATCATCTGTGGGCGAGGAGACACC 726
Qy 241 ArgArgIleGlyAspTyrLysValLysTyrGlyTyrThrAspIleAspLeuLeuSerAla 260
Db 727 CGCGGATCGGGGATTACAAGTTAAATATGGCTACACGACATTCACCTTCTCAGCGCT 786
Qy 261 AlalysSerLysProIleIleAlaGluProGluIleHisGlyAlaGlnProLeuAspGly 280
Db 787 GCCAAGTCCAAACCAATCATCGCAGAGCCAGAAATCCATGGGCGACAGCCGCTGGATGGG 846
Qy 281 ValThrGlyPheLeuValLeuMetSerGluGlyLeuTyrLysAlaLeuAlaHis 300
Db 847 GTGACGGGCTTCTTGGTCTGATGTGGAGGGGTTGTACAAAGGCCCTTAGAGGCGAGCCAT 906
Qy 301 GlyProGlyGlnAlaAsnGlnGluIleAlaMetIleAspThrGluPheAlaLysGln 320
Db 907 GGGCTGGCGAGCCAAACAGGAGATGTGTCGATGATTGACACTCAGTTTGCACAGCAG 966
Qy 321 ThrSerLeuAspAlaValAlaGlnAlaValValAspArgValLysArgIleHisSerAsp 340
Db 967 ACCTCCCTGCAGCAGTGGCCCGCCAGGCCGCTGTGGAGCCGGTGAAGCGCATCCACAGCGAC 1026
Qy 341 ThrPheAlaSerGlyGlyGluArgAlaAtgPheCysProArgHisGluAspMetThrLeu 360
Db 1027 ACCTTCGCCAGTGGTGGGAGCGTGCAGGGTTCTGCCCGGCGCAGGACATGACCCCTG 1086
Qy 361 LeuValArgAsnPheGlyTyrProLeuGlyGluMetSerGlnProThrProSerProAla 380
Db 1087 CTAGTGAGGAACCTTGGCTACCCGCTGGCGAAATGAGCCAGGCCCAACCGAGCCCGACGCC 1146
Qy 381 ProAlaAlaGlyGlyArgValTyrProValSerValProTyrSerSerAlaGlnSerThr 400
Db 1147 CCAGCTCCAGGAGGAGAGTGTACCTGTGTCTGTGCCATCTCCAGCGCCCGAGACACC 1206
Qy 401 SerLysThrSerValThrLeuSerLeuValMetProSerGlnGlyGlnMetValAsnGly 420
Db 1207 AGCAAGACCAAGCGTGACCCCTCTCCCTTGTCTATGCCCTCCAGGGCCAGATGGTCAACGGG 1266

Qy 421 AlaHisSerAlaSerThrLeuAspGluAlaThrProThrLeuThrAsnGlnSerProThr 440
Db 1267 GCTACAGTGTCTTCACCCCTGGAGAGAACCCCACTTCACCAACCAACCCGAGCC 1326
Qy 441 LeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSerAspGly 460
Db 1327 TTAACTGTGACGTCCACCAACAGCAGCAGCAGCAGCTCCAGCTCTGACGAGGC 1386
Qy 461 LeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGlyAlaGluPro 480
Db 1387 CTCTTCGGCTCCCGCGCCGCGCCACTCCCTCCGCTGCGAGAGCGGTCTGTGTAGCCC 1446
Qy 481 TyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGlyGluGlnSerVal 500
Db 1447 TATGTGACCTTGTGTAGTTTACCGCTTGAGCGTGAACCACTGGCGAGCAGCGCTG 1506
Qy 501 ValThrAlaPro 504
Db 1507 GTGACAGCACCG 1518

RESULT 6
US-09-529-279-42
; Sequence 42, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHITAKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529, 279
; CURRENT FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 42
; LENGTH: 1568
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11) .. (1549)
; US-09-529-279-42

Alignment Scores:
Pred. No.: 9.72e-249 Length: 1568
Score: 2577.00 Matches: 503
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.80% Mismatches: 0
Query Match: 99.88% Indels: 0
Gaps: 0

US-09-830-144-4 (1-504) x US-09-529-279-42 (1-1568)

Qy 1 MetAlaAlaGlnArgSerLeuGlnSerGluGlnGlnProSerTyrThrAsp 20
Db 38 ATGGCGGCGCAGAGAGAGAGCTTGCTGACAGATGAGCAGCAGCCAACTGGACAGATAC 97
Qy 21 LeuProLeuGlyHisLeuSerGlyValGlySerAlaSerAsnArgSerTyrSerAla 40
Db 98 CTGCTCTCTGACCTCTCTGAGGTGAGCTCAAGCTCCAAACCGAGCTACTCTGCTAAT 157
Qy 41 GlyIysGlyThrGlnSerHisProGluAspSerTyrPheLeuPheArgSerGluAsn 60
Db 158 GCGAAGGAGCAGAGAGCAGCAGCCGCGCAGAGAGCAGCTGCTCAAGTTGAGAGTGAAC 217
Qy 61 AsnGlyPheLeuTyrGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsnHeVal 80
Db 218 AACTGCTCTCTGTATGGGGCTTTCAACGGCTATGATGAGCAACCGAGTACCAGCTTCGTTG 277

Qy 81 AlaGlnArgLeuSerAlaGluLeuLeuGlyGlnLeuAsnAlaGluHisAlaGluAla 100
Db 278 GCCCAGCGGCTGTGCTGCGAGAGACTCTGCTGGCCAGCTGAATGCCAGCAGCGAGGCC 337
Qy 101 AspValArgValValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGlnSer 120
Db 338 GATGTGCGCGGTGTGCTCTGAGGCTTCGATGTGTGGAGAGAGAGCTTCTGTGAGTCC 397
Qy 121 IleAspAspAlaLeuAlaGluLysAlaSerLeuGlnSerGlnLeuProGluGlyValPro 140
Db 398 ATTTGACGAGCTTGTGCTGAGAGAGCAAGCTTCAGTGCAGATTTGCCAGAGGAGTCCCT 457
Qy 141 GlnHisGlnLeuProGluGlnLysIleLeuGlnArgLeuValArgLeuValArg 160
Db 458 CAGCAGCAGCTGCTCTCCAGATCAGAAAGATCCTTGAGAGCTCAAGACCTTGAGAGAG 517
Qy 161 GluIleSerGlyGlyAlaMetAlaValAlaValLeuLeuAsnAsnLysLeuTyrVal 180
Db 518 GAAATTTCCGGAGGGGCCATGGCTGTGTGGCGGTCTTCAACCAACCTCTACGTC 577
Qy 181 AlaAsnValGlyThrAsnArgAlaLeuLeuGlnLysSerThrValAspGlyLeuGlnVal 200
Db 578 GCCAATGTGATCAACCCGTGCACTTTTATGCAAAATGACAGTGGATGGTTGGAGGTG 637
Qy 201 ThrGlnLeuAsnValAspHisThrThrGluAsnGluAspGluLeuPheArgLeuSerGln 220
Db 638 ACACAGCTGAAGTGGACCAACCAACAGAGAGAGATGAGCTCTTCGCTTCGCGAG 697
Qy 221 LeuGlyLeuAspAlaGlyLysIleLysGlnValGlyIleLysGlyGlnGlnSerThr 240
Db 698 CTGGCTGTGATGTGAGAAAGATCAAGCAGTGGGAGATCATGTGGCGCAGAGAGCAC 757
Qy 241 ArgArgIleGlyAspTyrLysValLysTyrGlyTyrThrAspIleAspLeuSerAla 260
Db 758 CGCGGATCGGGATTCAGAGTTAAATGTGCTACAGGAGATTTGACCTTCCAGCGGT 817
Qy 261 AlaLysSerLysProIleIleAlaGluProGluIleHisGlyAlaGlnProLeuAspGly 280
Db 818 GCCAATCCAAACCAATCATCCGAGAGCAGAAATCCATGGGGCAGCGCTGATGGG 877
Qy 281 ValThrGlyPheLeuValLeuMetSerGluGlyLeuTyrLysAlaLeuGlnAlaHis 300
Db 878 GTGACGGGCTTCTGTGTGTGATGTGAGGGGTGTGTAAAGGCCCTTAAGGACAGCCAT 937
Qy 301 GlyProGlyGlnAlaAsnGlnIleAlaMetIleAspThrGluPheAlaLysGln 320
Db 938 GGGCTGGGAGGCCCAACAGAGATTTGCTGCGATGATGACACTGAGTTGCCAAGCAG 997
Qy 321 ThrSerLeuAspAlaValAlaGlnAlaValAlaAspArgValLysArgIleHisSerAsp 340
Db 998 ACCTCCCTGGACGCAATGGCGCCAGCGTCGTGAGACCGGAGGAAGGCAATCCACAGCGAC 1057
Qy 341 ThrPheAlaSerGlyGlyLysArgAlaArgPheCysProArgHisGluAspMetThrLeu 360
Db 1058 ACCCTGCGCAGTGTGGGAGCGTGCAGATTCGTGCCCCGGCAGCAGAGCAATGACCTTG 1117
Qy 361 LeuValArgAsnPheGlyTyrProLeuGlyGluMetSerGlnProThrProSerProAla 380
Db 1118 CTAGTAGAGAACTTTGGCTTACCGCTGGGCCAAATAGCCAGCGCCACACGAGCCAGCC 1177
Qy 381 ProAlaAlaGlyLysArgValTyrProValSerValProTyrSerSerAlaGlnSerThr 400
Db 1178 CCAAGTCAGAGAGAGCAGAGTACCTGTGTGTGCTCCATTAATCCAGCCGCCAGAGACACC 1237
Qy 401 SerLysThrSerValThrLeuSerLeuValMetProSerGlnGlyGlnMetValAsnGly 420
Db 1238 AGCAAGACCAAGCGTGAACCTTCTCTGTATGCTCTCCAGGGGCGAGATGTCAAGGAG 1297
Qy 421 AlaHisSerAlaSerThrLeuAspGluAlaThrProThrLeuThrAsnGlnSerProThr 440
Db 1298 GCTACAGTGTCTTCCACCTCTGAGAGAAAGCACCACCTTACCAACCAAGCCCGAGCC 1357

QY 341 ThrPheAlaSerGIyGIyGluArgAlaArgPheCysProArgHisGluAspMetThrLeu 360
Db 1050 ACCTTCCGCGAGTGGGGAGCGGTGCCAGGTTCTGCCCGGACGAGGACATGACCTTG 1109
QY 361 LeuValArgAsnPheGlyTyrProLeuGIyGluMetSerGlnProThrProSerProAla 380
Db 1110 CTAGTGAAGAACTTGGCTTACCCCTGGGCGAATGAGCCAGCCCAACCGAGCCGACGCC 1169
QY 381 ProAlaAlaGIyGIyArgValTyrProValSerValProTyrSerSerAlaGlnSerThr 400
Db 1170 CCAGCTGAGAGAGGAGGAGGAGTACCTGTCTGTGCATCTCAGCGCCAGAGCACC 1229
QY 401 SerTyrThrSerValThrLeuSerLeuValMetProSerGlnGlyGluMetValAsnGly 420
Db 1230 AGCAAGACGACGTCGACCTCTCCCTTGTCTATGCTCCCGAGGGCGAGATGTCAACGG 1289
QY 421 AlaHisSerAlaSerThrLeuAspGluAlaThrProThrLeuThrAsnGlnSerProThr 440
Db 1290 GCTCACAGTCTTCCACCTTGAGAGAGCAGCACCCTCACCACCAAGCCGAGACC 1349
QY 441 LeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSerSerAspGlyGly 460
Db 1350 TTAACCTCGACGTCCACCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1409
QY 461 LeuPheAlaSerArgProAlaHisSerLeuProProGIyGIyAspGlyArgValGluPro 480
Db 1410 CTCTTCCGCTCCCGGCGCCGCTGCTCCGCTGGCGAGAGAGTGTGTTGAGGCC 1469
QY 481 TyrValAspPheAlaGluPheTyrArgLeuTyrSerValAspHisGlyGluGlnSerVal 500
Db 1470 TATGTGAGACTTGTCTGAGTTTATACGCTCTGAGCGTGGAGACATGGGAGAGAGCCTG 1529
QY 501 ValThrAlaPro 504
Db 1530 GTGACAGCACCG 1541

RESULT 8
US-09-144-178-5
Sequence 5, Application US/09144178
Patent No. 5989862
GENERAL INFORMATION:
APPLICANT: MATSUMOTO, Kunihiko
APPLICANT: NISHIDA, Eisuke
TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/144,178
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/752,891
FILING DATE: 20-NOV-1996
APPLICATION NUMBER: JP 8-300856
FILING DATE: 28-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 8-126282
FILING DATE: 24-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768

REFERENCE/DOCKET NUMBER: 17981/111
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5339
TELEX: 904136
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1560 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 30..1541
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 30..1541
US-09-144-178-5

Alignment Scores:
Pred. No.: 1,53e-248 Length: 1560
Score: 2575.00 Matches: 503
Percent Similarity: 99.80% Conservative: 0
Best Local Similarity: 99.80% Mismatches: 1
Query Match: 99.81% Indels: 0
Gaps: 0

US-09-830-144-4 (1-504) X US-09-144-178-5 (1-1560)

QY 1 MetAlaAlaGlnArgSerLeuLeuGlnSerGlnGlnProSerThrThrAsp 20
Db 30 ATGGCGCGGAGAGGAGGAGGAGCTTGTGCAGAGTGAACAGACAGCAAGCTTGACAGATGAC 89
QY 21 LeuProLeuCysHisLeuSerGIyValGlySerAlaSerAsnArgSerTyrSerAlaAsp 40
Db 90 CTGCTCTTCTGCCACCTCTTGTGGGTTGGCTCAGCTCCACCGCAGCTACTCTGCTGAT 149
QY 41 GlyLysGIyThrGlnSerHisProProGIyAspSerTyrLeuLysPheArgSerGluAsn 60
Db 150 GGCAAGGACACTGAGAGCCACCGCCAGAGACAGATGGCTCAAGTTCAAGACTGAGAAC 209
QY 61 AsnCysPheLeuTyrGIyValPheAsnGIyTyrAspGIyAsnArgValThrAsnPheVal 80
Db 210 AACGCTTCTCTGATGGGCTTTCACGGCTATGATGCAACCGATGACCACTTCGTCG 269
QY 81 AlaGlnArgLeuSerAlaGluLeuLeuGlnGlnLeuAsnAlaGlnHisAlaGluAla 100
Db 270 GCCAAGCGGCTGTCCGAGAGCTCTGCTGGCCAGCTGAATGCCAGCCGAGAGCC 329
QY 101 AspValArgArgValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGlnSer 120
Db 330 GATGTGGCGCGTGTGCTGCTGACAGCCCTTGATGTGTGAGAGAGACTTCTGAGATCC 389
QY 121 IleAspAspAlaLeuAlaGluValAspLeuGlnSerGlnSerGlnSerProGIyValPro 140
Db 390 ATTGACAGCGCTTGGCTGAGAGGCAAGCTCCAGTCTGCAATTTGCCAGAGGAGTCCCT 449
QY 141 GlnHisGlnLeuProProGIyTyrGlnLysIleLeuGluArgLeuLysThrLeuGluArg 160
Db 450 CAGCACGACGCTGCTCCCTCACTATCAGAAATCTCTGAGAGACTCAAGAGCTTGAAGAGG 509
QY 161 GluIleSerGIyGIyAlaMetAlaValValAlaLeuLeuAsnLeuLysLeuTyrVal 180
Db 510 GAAATTTGGGAGGGGCGCATGGCGGTGTCTTCAACACACAGCTTACGTC 569
QY 181 AlaAsnValGIyThrAsnArgAlaLeuLeuCysLysSerThrValAspGIyLeuGlnVal 200
Db 570 GCCAATGCTCGTACCAACCGTCACTTTATGCAATGACAGTGGATGGGTTCAAGTGC 629
QY 201 ThrGlnLeuAsnValAspHisThrThrGluAsnGluAspGIyLeuPheArgLeuSerGln 220
Db 630 ACACAGCTGAAGTGAACACACACACAGAGAGATGAGTCTTCCGCTTTCGAG 689

QY 221 LeuGlyLeuAspAlaGlyLysIleLysGlnValGlyIleIleCysGlyGlnGluSerThr 240
Db 690 CTGGGCTTGGATCTGGAAAGATCAAGCAGGTGGGGATCATCTGTGGCGAGAGCACC 749
QY 241 ArgArgIleGlyAspTyrLysValLysTyrGlyTyrThrAspIleAspLeuLeuSerAla 260
Db 750 CGCGGATCGGGGATTACAAGGTTAAATATGCTACAGGACATTGACCTTCTCAGGCT 809
QY 261 AlaLysSerLysProIleIleAlaGluProIleIleHisGlyAlaGlnProLeuAspGly 280
Db 810 GCCAAGTCCAAACCAATCATCGAGAGCAGAAATCCATGGGCACAGCCCTGGATGG 869
QY 281 ValThrGlyPheLeuValLeuMetSerGluGlyLeuTyrLysAlaLeuGluAlaHis 300
Db 870 GTGAGGGCTTCTGGTGTGATGTCGAGGGGTGTACAGGCCCTAGAGGAGCCCAT 929
QY 301 GlyProGlyGlnAlaAsnGlnGluIleAlaAlaMetIleAspThrGluPheAlaLysGln 320
Db 930 GGGCTGGCAGGCCAACCCAGGAGATTGCTCGGATGATTGACACTGAGTTTCCAGCAG 989
QY 321 ThrSerLeuAspAlaValAlaGlnAlaValAlaAspArgValLysArgIleHisSerAsp 340
Db 990 ACCTCCCTGGACGAGTGGCCAGGCCGCTGTCGACCGGTGAAGCGCATCCACAGCGAC 1049
QY 341 ThrPheAlaSerGlyGlyGluArgAlaArgPheCysProArgHisGluAspMetThrLeu 360
Db 1050 ACCTTCGCCAGTGTGGGAGCGTGCAGGTTCTGCCCCCGGCACGAGACATGACCTTG 1109
QY 361 LeuValArgAsnPheGlyTyrProLeuGlyGluMetSerGlnProThrProSerProAla 380
Db 1110 CTAGTGAGGAACCTTGGCTACCCGCTGGCGGAAATGAGCCACACCGAGGCCAGCC 1169
QY 381 ProAlaAlaGlyArgValTyrProValSerValProValSerTyrSerSerAlaGlnSerThr 400
Db 1170 CCAGCTCGAGAGACAGGTGACCTGTGTGTGTCGATCTCCAGCGCCCGAGAGCACC 1229
QY 401 SerLysThrSerValThrLeuSerLeuValMetProSerGlnGlyMetValAsnGly 420
Db 1230 AGCAAGACCGAGTGCCTCTCCCTGTGTCAGTCCCTCCAGGCCAGATGTCACCGGG 1289
QY 421 AlaHisSerAlaSerThrLeuAspGluAlaThrProThrLeuThrAsnGlnSerProThr 440
Db 1290 GCTCACAGTCTCCACCTTGGACGAAGCCACCCGCCCTCCACCAACCAAGCCGACC 1349
QY 441 LeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSerSerSerSerSer 460
Db 1350 TTAACCTCGAGTCCACCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1409
QY 461 LeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGlyArgValGluPro 480
Db 1410 CTCTTCGCTCCCGGCCCGCCACTCGCTCCCGCTGGCGAGGACGTCGTGTTGAGGCC 1469
QY 481 TyrValAspPheAlaGluPheTyrArgLeuTyrSerValAspHisGlyGluGlnSerVal 500
Db 1470 TATGTGACTTTGCTGAGTTTTACCGCCTCTGGAGCGTGGACCATGGCGAGCAGCGGT 1529
QY 501 ValThrAlaPro 504
Db 1530 GTGACAGCACC 1541

RESULT 9

US-09-406-854-5

; Sequence 5, Application US/09406854

; Patent No. 6140042

; GENERAL INFORMATION:

; APPLICANT: MATSUMOTO, Kunihiro

; APPLICANT: NISHIDA, Eisuke

; TITLE OF INVENTION: TABI PROTEIN AND DNA CODING THEREFOR

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Foley & Lardner

; STREET: 3000 K Street, N.W., Suite 500

; CITY: Washington

; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/406,854
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/752,891
; FILING DATE: 20-NOV-1996
; APPLICATION NUMBER: JP 8-300856
; FILING DATE: 28-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-126282
; FILING DATE: 24-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 17981/111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1560 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 30..1541
; NAME/KEY: mat_peptide
; LOCATION: 30..1541
; US-09-406-854-5
Alignment Scores:
Pred. No.: 1.53e-248 Length: 1560
Score: 2575.00 Matches: 503
Percent Similarity: 99.80% Conservative: 0
Best Local Similarity: 99.80% Mismatches: 1
Query Match: 99.81% Indels: 0
DB: 3 Gaps: 0
US-09-830-144-4 (1-504) x US-09-406-854-5 (1-1560)
QY 1 MetAlaAlaGlnArgArgSerLeuLeuGlnSerGluGlnGlnProSerTrpThrAspAsp 20
Db 30 ATGGCGGCGCAGAGGAGGAGCTTCTGCAGAGTGAGCAGCAGCAGCAGCAGCAGCAGTGC 89
QY 21 LeuProLeuCysHisLeuSerGlyValGlySerAlaSerAsnArgSerTyrSerAlaAsp 40
Db 90 CTGCCTCTCTGCCACCTCTCTGGGGTTGGCTCAGCCTCCCAACCGCAGCTACTCTGCTGAT 149
QY 41 GlyLysGlyThrGluSerHisProGluAspSerTrpLeuLysPheArgSerGluAsn 60
Db 150 GGCAGGGCAGCTGAGAGCCACCCCGCAGAGACAGATGGCTCAAGTTCAGGAGTGAGAAC 209
QY 61 AsnCysPheLeuTyrGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsnPheVal 80
Db 210 AACTGCTTCTGTATGGGTCTTCAACGGCTATGATGGCAACCGAGTGACCAACTTCGTG 269
QY 81 AlaGlnArgLeuSerAlaGluLeuLeuGlyGlnLeuAsnAlaGluHisAlaGluAla 100
Db 270 GCCACAGCGGCTGTCGCGAGAGCTCTCTGCGGCAGCTGAATGCCGACGACGCCGAGGCC 329
QY 101 AspValArgArgValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGluSer 120

Db 302 GCAAGAGAAAACC-----TCCGAGGAAGAGAAGATGGCAGTGAAGAGCTTGTG 352
Qy 22 -----ProLeuCysHisLeuSerGlyValGlySerAlaSerAsnArgSerTyrSerAla 39
Db 353 GAAAGAAAGTTTGAAGCCTCTTCGGTG---ATCTTTGGTCTGAAGGGCTATGTGCT 409
Qy 40 AspGlyValGly-----Thr 44
Db 410 GAGCGGAAGGGTGAGAGGAGAGAGATGCCAGGTATCCCTCATCTCAACGACATCAC 469
Qy 45 GluSerHisProProGluAspSerTrpLeuLysPheArgSerGluAsnAsnCysPheLeu 64
Db 470 GAGGAGTGTAGCCCCCTTCCTCAT-----ACTCGGTTTTCATAT 514
Qy 65 TyrGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsnPheValAlaGlnArgLeu 84
Db 515 TTTGCTGTTTTTGTGACATGAGGAATTCGAGCCTCAAAATTTGCTGCACAGAAATTG 574
Qy 85 SerAlaGluLeu-----LeuGlyGlnLeuAsnAlaGluHisAlaGluAlaAspVal 102
Db 575 CATCAAACTTAATCAGAAAATTTCTCAAGGAGATGTAATCAGTGTAGAGAAAACCGTG 634
Qy 103 ArgArgValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGluSerIleAsp 122
Db 635 AAGAGATGCTTTTGGACACTTTCAGCATACTGATGAGAGTTCCT-----682
Qy 123 AspAlaLeuAlaGluLysAlaSerLeuGlnSerGlnLeuProGluGlyValProGlnHis 142
Db 683 -----AAACAAGCTTCAGCCCAAG-----703
Qy 143 GlnLeuProProGlnTyrGlnLysIleLeuGluArgLeuLysThrLeuGluArgGluIle 162
Db 704 -----CCTGCCCTGGAAA-----715
Qy 163 SerGlyGlyAlaMetAlaValAlaValAlaLeuLeuAsnAsnLysLeuTyrValAlaAsn 182
Db 716 ---GATGGTCCATCCACGCTGTCTTGGCTGTAGACAACTTTATATTGCCAAC 772
Qy 183 ValGlyThrAsnArgAlaLeuLeuCysLysSerThrValAspGlyLeuGln-----Val 200
Db 773 CTCGGAGATAGTCGGCAATCTTGTCTGTATATGAGGAGAGTCAAAACATCGAGCC 832
Qy 201 ThrGlnLeuAsnValAspHisThrThrGluAsnGluAspGluLeuPheArgLeuSerGln 220
Db 833 TTAAGCCTCAGCAAGAGCATATCAACTCAGTATGAAGCGGATGAGGATACAGAG 892
Qy 221 LeuGlyLeuAspAlaGlyLysIleLysGlnValGlyIleIleCysGlyGlnGluSerThr 240
Db 893 CTGGA-----CGAAACGTGAGGATGGCGTGTTCGGCGTGTAGAGGTGTCA 943
Qy 241 ArgArgIleGlyAsp-----TyrLysValLysTyrGlyThrAspIleAspLeuLeu 258
Db 944 CGCTCCATTGGGACGGCAGTACAAG---CGTGGCGGTGCACCTCTGT-----991
Qy 259 SerAlaAlaLysSerLysProIleIleAlaGluProGluIleHisGlyAlaGln-----276
Db 992 -----CCGACATCAGACGCTGCCAGCTGACC 1018
Qy 277 ProLeuAspGlyValThrGlyPheLeuValLeuMetSerGluGlyLeuTyrLysAlaLeu 296
Db 1019 CCAATGAC-----AGGTTCAATTTGTTGGCCTGTGATGGGCTCTTCAAGGCTTT 1069
Qy 297 GluAlaAlaHisGlyProGlyGlnAlaAsnGlnGluIleAlaAlaMetIleAspThrGlu 316
Db 1070 ACC-----CCAGAAGAAGCGGTGAACCTTCACTTCTGCTCTCGAGGATGAA 1117
Qy 317 PheAlaLys-----GlnThrSerLeuAspAlaValAlaGlnAlaValValAsp 332
Db 1118 AAGATCCAGCCCGGAAGGAAGTCGACAGCCGCGCTACGAAGCAGCCTGCAAC 1177
Qy 333 ArgVal 334
Db 1178 AGGCTG 1183

RESULT 11
US-08-764-233A-1
; Sequence 1, Application US/08764233A
; Patent No. 5716849
; GENERAL INFORMATION:
; APPLICANT: Ligon, James M.
; APPLICANT: Schupp, Thomas
; APPLICANT: Beck, James J.
; APPLICANT: Hill, Dwight S.
; APPLICANT: Neff, Snezana
; APPLICANT: Ryals, John A.
; TITLE OF INVENTION: Genes For The Biosynthesis Of Soraphen
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: 520 White Plains Road, P.O. Box 2005
; CITY: Tarrytown
; STATE: NY
; COUNTRY: USA
; ZIP: 10591
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/764,233A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/729,214
; FILING DATE: 09-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/258,261
; FILING DATE: 08-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: 1506/CIP6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919) 541-8587
; TELEFAX: (919) 541-8689
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 49377 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Sorangium cellulosum
; IMMEDIATE SOURCE:
; CLONE: p98/1, pJL3, and pVKM15
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 383..760 /product= "SorR"
; OTHER INFORMATION: /note= "This gene encodes a protein that is highly homologous
; OTHER INFORMATION: the reductase domains of type I PKs such as eryA from
; OTHER INFORMATION: Saccharopolyspora erythraea."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 927..19874 /product= "SorA"
; OTHER INFORMATION: /note= "Gene product is highly homologous to type I PKs that
; OTHER INFORMATION: are known to be involved in the synthesis of polyketide
; OTHER INFORMATION: compounds."
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 942..7115 /product= "Module 1 of SorA"
; OTHER INFORMATION: /product= "Module 1 of SorA"
; FEATURE:

```

/ NAME/KEY: misc_feature
/ LOCATION: 7203..12884
/ OTHER INFORMATION: /product= "Module 2 of SorA"
/
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 13455..19616
/ OTHER INFORMATION: /product= "Module 3 of SorA"
/
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 19871..46318
/ OTHER INFORMATION: /product= "SorB"
/
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 19870..24556
/ OTHER INFORMATION: /product= "Module 1 of SorB"
/
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 24638..30820
/ OTHER INFORMATION: /product= "Module 2 of SorB"
/
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 30881..35446
/ OTHER INFORMATION: /product= "Module 3 of SorB"
/
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 35528..40114
/ OTHER INFORMATION: /product= "Module 4 of SorB"
/
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 40190..46318
/ OTHER INFORMATION: /product= "Module 5 of SorB"
/
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 46851..47891
/ OTHER INFORMATION: /product= "SorM"
/
/ OTHER INFORMATION: /note= "The protein encoded by the sorM gene is highly
/ OTHER INFORMATION: homologous to the methyltransferase from Streptomyces
/ OTHER INFORMATION: hygroscopicus that is involved in the synthesis of the
/ OTHER INFORMATION: polyketide rapamycin."
/
US-08-764-233A-1

Alignment Scores:
Pred. No.: 0.497
Score: 131.00
Percent Similarity: 33.56%
Best Local Similarity: 22.76%
Query Match: 5.08%
DB: 1
Length: 49377
Matches: 99
Conservative: 47
Mismatch: 146
Indels: 143
Gaps: 19

US-09-830-144-4 (1-504) x US-08-764-233A-1 (1-49377)
QY 84 LeuSerAlaGluLeuLeuGluGlnLeuAsnAlaGluHisAlaGluAlaAspValArg 103
Db 9057 CTGAAGCCCGCTCTCTCTCGGACACATCCATGCGCGAGCTGTCGCCGCCACGTCGCC 9116
QY 104 ArgValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGluSerIleAspAsp 123
Db 9117 GCGGCTCTT-----TCTCTCCAGGAC 9137
QY 124 Ala-----LeuAlaGluLysAlaSerLeuGlnSerGlnLeuProGluGlyValPro 140
Db 9138 GCGTCACCCCTGTCGCGCGCCCAAGCTCATGCAAGGCGTCCCAAGGCGCGCC 9197
QY 141 GlnHisGlnLeuProGlnGlnTyrGlnLysIleLeuGluArgLeuValThrLeuGluArg 160
Db 9198 ATGTCACCTTCGAGCTCCGAGAGAGAGATCGCGACCTTCTCCAGCCCTACGATGA 9257
QY 161 GluIleSerGlyGlyAlaMetAlaValValAlaValLeuLeuAsnLysLeuTyrVal 180
Db 9258 CGAGTACCTTCGCGCC-----CTCAATGGGCTCTCTCCACC 9296
QY 181 AlaAsnValGlyThrAsnArgAlaLeuLeu-----CysLysSerThrValAspGly 197

```

```

Db 9297 GTCTGCTGCGCATGAAGACGCGGTGTGGATGCGCCCGACGCGCAAGCCCTCGGA 9356
QY 198 LeuGlnValThrGlnLeuAsnValAspHisThrThrGluAsnGlu----- 212
Db 9357 CGAAGACACACAGCTGCGCGGTGAGGACGCTTCACCTCCGACATGACGGAATG 9416
QY 213 --AspGluLeuPheArgLeuSerGln--LeuGlyLeuAspAlaGlyIleLysGln 230
Db 9417 CTCGACGACCTTCGCGCGGTGCGCCCAAGCTTACCTACCATCCCGACGATCCCATC 9476
QY 231 ValGlyIleIleCysGlyGlnGluSerThrArgArgIleGlyAspTyrLysValLysTyr 250
Db 9477 ATCTCAACGTACCGCGCGCGCGCCACG-----GACCAACGAG----- 9515
QY 251 GlyTyrThrAspIleAspLeuLeuSerAlaAlaLysSerLysProIleIleAlaGluPro 270
Db 9516 -----CTGCGCTGCGCC 9527
QY 271 GluIle-----HisGlyAlaGlnProLeuAspGlyValThrGlyPhe 284
Db 9528 GACTACTGGGTCCGCCACGCTTCCGACACCGCTCCGCTTCTCCACGCGGTA----- 9578
QY 285 LeuValLeuMetSerGluGlyLeuTyrLysAlaLeuGluAlaAlaHisGlyProGlyGln 304
Db 9579 -----CGTGCCTTCACCGCGGAAGG--- 9599
QY 305 AlaAsnGlnGluIleAlaAlaMetIleAspThrGluPheAlaLysGlnThrSerLeuAsp 324
Db 9600 -----GCACGCGCTTTCTCGACCTCGGCGCTCACGCTGCTCTCTCC 9641
QY 325 AlaValAlaGlnAlaValAlaVal--AspArgValLys--ArgIleHisSerAspThrPheAl 343
Db 9642 GCGCTTCGCAAGACGCGCTTCGACACAGACGAAAGACGTCGCAATCGGCTTCTTCC 9701
QY 343 AserGlyGlyGluArgAla-----ArgPheCysProArgHisGluAspMetThr 359
Db 9702 ACCCTCCGCAAGGACGCGCGACGCGCGGAGCGTTCACGCGCGCGCTGCGCTCCAC 9761
QY 359 fLeuLeuValArgAsnPheLysTyrProLeuGlyGluMetSerGln----- 374
Db 9762 GCTG-----CAGGCTCACACCTGAGCTGAGGCTTCTTCGCCCCCTTCGCTCCA 9812
QY 375 -----ProThrProSerProAlaProAlaAlaGlyLys----- 385
Db 9813 CGCAAGGTCTCCCTCCCACTATGCTTCCAGCGGAGCGCTTCTGCTCGATGCTTCC 9872
QY 386 ----ArgValTyrProValSerValProTyrSerSerAlaGlnSerThrLysThrSe 404
Db 9873 AAGCACACGCTGCGACGTCGCTCCGACGCGCTGACCTCGACCGATACCCGCGTCTC 9932
QY 404 fValThrLeuSerLeuValMetProSerGlnGlnMetValAsnGlyAlaHisSerAl 424
Db 9933 GGGCGCGCGCTCC----- 9945
QY 424 AserThrLeuAspGluAlaThrProThrLeuThrAsnGlnSerProThrLeuThrLeuGl 444
Db 9946 -----CCCTCGCGGACCGCATGGCTTC 9968
QY 444 nSerThr-----AsnThrHisThrGlnSer-----SerSerSerSe 457
Db 9969 CTCTTACAGACGACTCTCAGCTCAGAGATCCGTGCTGCGCGATACGCTGCTTC 10028
QY 457 fAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuPro 471
Db 10029 GGTACACCATCTTCCGGGACCTGCTTTCGAGCTTGGCC 10071

RESULT 12
US-08-822-701-1
; Sequence 1, Application us/08822701
; Patent No. 5976853
; GENERAL INFORMATION:
; APPLICANT: Guthridge, Mark

```

```

/ APPLICANT: Basilio, Claudio
/ TITLE OF INVENTION: NOVEL GROWTH FACTOR INDUCIBLE
/ TITLE OF INVENTION: SERINE/THREONINE PHOSPHATASE, FIN13
/ NUMBER OF SEQUENCES: 18
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: David A. Jackson, Esq.
/ STREET: 411 Hackensack Ave, Continental Plaza, 4th
/ STREET: Floor
/ CITY: Hackensack
/ STATE: New Jersey
/ COUNTRY: USA
/ ZIP: 07601
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/822,701
/ FILING DATE:
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Jackson Esq., David A.
/ REGISTRATION NUMBER: 26,742
/ REFERENCE/DOCKET NUMBER: 1049-1-002 N
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 201-487-5800
/ TELEFAX: 201-343-1684
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1824 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ US-08-822-701-1
/
/ Alignment Scores:
/ Pred. No.: 0.0039
/ Score: 128.50
/ Percent Similarity: 36.60%
/ Best Local Similarity: 21.61%
/ Query Match: 4.98%
/ DB: 2
/
/ US-08-820-144-4 (1-504) x US-08-822-701-1 (1-1824)
/ Length: 1824
/ Matches: 75
/ Conservative: 52
/ Mismatches: 129
/ Indels: 91
/ Gaps: 12

```

FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 1049-1-002 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1824 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHEICAL: NO
ANTI-SENSE: NO
US-08-935-855-1

Alignment Scores:
Pred. No.: 0.0039 Length: 1824
Score: 128.50 Matches: 75
Percent Similarity: 36.60% Conservative: 52
Best Local Similarity: 21.61% Mismatches: 129
Query Match: 4.98% Indels: 91
DB: 3 Gaps: 12

US-09-830-144-4 (1-504) x US-08-935-855-1 (1-1824)

QY 165 GAlaMetAlaValValAlaValLeuLeuAsnAsnLeuTyValAlaValAlaVal 184
DB 736 GGCACCAACAGCGGTGGCTGTGATCAGAGGAGCAAGCATTTATGGCCATGACAGA 795
QY 185 ThrAsnArgAlaLeuLeuCysLysSerThrValAspGluValThrGlnLeuAsn 204
DB 796 GACTCTCGCTGTGTGTGTCCGAGGCTGGC-----AAAGCTTTAGATATGTCC 843
QY 205 ValAspHisThrThrGluAsnGluAspGluLeuPheArgLeuSerGlnLeuGluAsp 224
DB 844 TATGACCAACCAACGAGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 897
QY 225 AlaGlyLysIleLysGlnValGly---IleIleCysGlyGlnGlnSerThrArgArgIle 243
DB 898 ---GGCAAGGTCACCATGATGATGATGATGATGATGATGATGATGATGATGATG 954
QY 244 GlyAspTyrLysValLysTyrGlyTyrThrAspIleAspLeuSerAlaIleLysSer 263
DB 955 GGAGCCACCTTCTCAAG-----AGAAAC 978
QY 264 LysProIleIleAlaGluProGluIleHisGlyAlaGlnPro----- 277
DB 979 AAAAATCTGCCACCCGAGGAGACAGATATTTCTGCCCTTCTGACATCAAGGTGCTG 1038
QY 278 LeuAspGluValThrGlyPheLeuValLeuMetSerGlnGlyLeuTyrLysAlaLeuGlu 1095
DB 1039 CTCCTGATGACCATGAATTCATGATGATGATGATGATGATGATGATGATGATG 1095
QY 298 AlaAlaHisGlyProGlyGlnAlaAsnGlnGluIleAlaAlaMetIleAspThrGluPhe 317
DB 1096 -----AGAGCCAGAGAGTGTGATGATTTATTCATTCATTCATTCATTCATTC 1134
QY 318 AlaLysGlnThrSer-----LeuAspAlaValAlaGlnAlaValAlaVal 331
DB 1135 AGTCAACGTGATGAAGAGGGAGCTTGGTTATTTGTCATTCATTCATTCATTCATTC 1194
QY 332 AspArg-----ValLysArgIle-His 339
DB 1195 GATCAGTGCCTGGGCGCAGACATTCCTGGGATGATGATGATGATGATGATGATGATG 1254
QY 339 ThrAspThrPheAlaSerGly-----G 346
DB 1255 ATCATCATTTGCTTCAAGCCCGAAGACAGTAGAGCTTCAAGCAGAGAGTGCAAGAGG 1314

QY 346 YGluArgAlaArgPheCysProArgHisGlnLysPheThrLeuLeuValArgAsnPheG 366
DB 1315 AAACGTGAGAGAGGACATGTCTCCACGAGAGGGGCTGAAGACACCGGCAACAGTGA 1374
QY 366 YTYrProLeuGlyGluMetSerGlnProThrPro-----SerPr 379
DB 1375 AAGCCAAAGAGGACATGATGATCAACCGAGCCCTGATGATGATGATGATGATGATG 1434
QY 379 O-AlaPro-----Ala 383
DB 1435 TTGACCCGAGACCTGAGTGTTCCTGCTTTCCTTTCCTTTCCTTTCCTTTCCTTTC 1494
QY 383 LaGlyGlyArgValTyrProValSerValProTyrSerSerAlaGlnSerThrSerLys 403
DB 1495 CAGGGGGCTGGTGTCTTCTTCACG---CAATTACAAAGAGGCCCCCCCCCCCCCA 1551
QY 403 hSerValThrLeuSerLeuValMetProSerGlnGlnMetValAsnGlyAlaHis 423
DB 1552 CGCGGACGCTGGAGGCTGCTGCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTC 1611
QY 423 eAlaSerThrLeuAspGluAlaThrProThrLeuThrAsnGlnSerProThrLeuThr 443
DB 1612 GACTATCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1671
QY 443 eugInSerThrAsnThr 448
DB 1672 TTTACTCTGTGACACT 1688

RESULT 14
US-08-935-855-21
Sequence 21, Application US/08935855
Patent No. 6066485
GENERAL INFORMATION:
APPLICANT: Guthridge, Mark
TITLE OF INVENTION: NOVEL GROWTH FACTOR INDUCIBLE
TITLE OF INVENTION: SERINE/THROMBIN PHOSPHATASE, FIN13
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: David A. Jackson, Esq.
STREET: 411 Hackensack Ave, Continental Plaza, 4th
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/935,855
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 1049-1-002 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 2081 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHEICAL: NO
ORGANISM: Mus musculus

US-08-935-855-21

Alignment Scores:

Pred. No.:	0.00484	Length:	2081
Score:	128.50	Matches:	75
Percent Similarity:	36.60%	Conservative:	52
Best Local Similarity:	21.61%	Mismatches:	129
Query Match:	4.98%	Indels:	91
DB:	3	Gaps:	12

US-09-830-144-4 (1-504) x US-08-935-855-21 (1-2081)

Qy	165	GlyAlaMetAlaValAlaValLeuLeuAsnAsnLysLeuTyrValAlaAsnValGly	184
Db	993	GGCACAACACGGTGGTCTGTATGATGAGGGAAGCAGTGTGTTGGCCAAATGCAGGA	1052
Qy	185	ThrAsnArgAlaLeuLeuCysLysSerThrValAspGlyLeuGlnValThrGlnLeuAsn	204
Db	1053	GACTCTCGCTGTGTGGTCCGAGGCTGGC-----AAAGCTTTAGATATGTC	1100
Qy	205	ValAspHisThrThrGluAsnGluAspGluLeuPheArgLeuSerGlnLeuGlyLeuAsp	224
Db	1101	TATGACCACAAACAGAGAGTGAAGTGGAGCTGGCAGCATCAAGAATGCTGGT-----	1154
Qy	225	AlaGlyLysIleGlyGlnValGly---IleIleCysGlyGlnGlnSerThrArgArgIle	243
Db	1155	---GGCAAGGTCAACATGGATGGAGAGTCAATGGAGGCTCAACCTCTCCAGGGCCATT	1211
Qy	244	GlyAspTyrLysValLysTyrGlyTyrThrAspIleAspLeuSerAlaAlaLysSer	263
Db	1212	GGAGACCACCTTCTACAG-----AGAAAC	1235
Qy	264	LysProIleAlaGluProGluIleHisGlyAlaGlnPro-----	277
Db	1236	AAAAAAGTTCGCCACCCCAACAGACAGATTTCTGCCCTTCTGACATCAAGGTGCTGACT	1295
Qy	278	LeuAspGlyValThrGlyPheLeuValLeuMetSerGluGlyLeuTyrLysAlaLeuGlu	297
Db	1296	CTCAGTGTGACCATGAATTCATGTCATCTGTGAGCGCATCGAATGTGATG---	1352
Qy	298	AlaAlaHisGlyProGlyClnAlaAsnGlnGluIleAlaAlaMetIleAspThrGluPhe	317
Db	1353	-----AGCAGCAGGAGGTGTGTAGACTTTTATTCAATCAAGATC	1391
Qy	318	AlaLysGlnThrSer-----LeuAspAlaValAlaGlnAlaValVal	331
Db	1392	AGTCACAGTGATGAACACGGGAGCTTCGGTTATTGTTCATCTTGGAGAGCTGCTG	1451
Qy	332	AspArg-----VallysArgIle-His	339
Db	1452	GATCAGTGTCTGGCGCCACAGACTTCTGGGATGTGTACAGGGTGTGACACATGACGTGC	1511
Qy	339	rAspThrPheAlaSerGly-----	346
Db	1512	ATCATCATTTGCTTCAAGCCCGAAACACAGTAGAGCTTCAGGCAGAGAGTGGCAAGG	1571
Qy	346	YGLuArgAlaArgPheCysProArgHisGluAspMetThrLeuLeuValArgAsnPheG	366
Db	1572	AAACTGGAGAGGCATGTCTCCAGCAGGGGGCTGAAGACACCGGCAACAGTGCACAAAAG	1631
Qy	366	YtyrProLeuGlyGluMetSerGlnProThrPro-----SerPr	379
Db	1632	AAGGCCAAGAGGAGCTAGTGTGTCAACCGGACCTGCCCATGTGGACTGTTTCTGAGCCC	1691
Qy	379	o-AlaPro-----Ala	383
Db	1692	TTGGACCCGAGACTGAGTTTGTCTTCTTGTAGCCTTAGCAGTGGGTATGAGGTGTG	1751
Qy	383	lnGlyArgValTyrProValSerValProTyrSerSerAlaGlnSerThrSerLysT	403
Db	1752	CAGGGGCTGGGTGCTTCTCTCAGC-----CCATTACAAAGAGGGCCCCCACCACCCCA	1808
Qy	403	hrSerValThrLeuSerLeuValMetProSerGlnGlyGlnMetValAsnGlyAlaHis	423

Db	1809	CGCGCAGCCTGGGAGGCTCTGTGTCTCTTAAGCCTCTTACTCTCTTGGGCTCATC	1868
Qy	423	erAlaSerThrLeuAspGluAlaThrProThrLeuThrAsnGlnSerProThrLeuThrL	443
Db	1869	GACTATCGGTTCTGTGCTGTGCTCTGTGTGTGGAGGAAGGAGGTAGTGTCTGATT	1928
Qy	443	eUGlnSerThrAsnThr	448
Db	1929	TTTACTCTGTGAACACT	1945

RESULT 15

US-08-258-261B-6

; Sequence 6, Application US/08258261B

; Patent No. 5639949

; GENERAL INFORMATION:

; APPLICANT: Schupp, Thomas

; APPLICANT: Ligon, James M.

; APPLICANT: Beck, James Joseph

; APPLICANT: Hill, Dwight Steven

; APPLICANT: Ryals, John Andrew

; APPLICANT: Gaffney, Thomas Deane

; APPLICANT: Lam, Stephen Ting

; APPLICANT: Hammer, Phillip E.

; APPLICANT: Uknes, Scott Joseph

; TITLE OF INVENTION: Genes for the synthesis of

; TITLE OF INVENTION: antipathogenic substances

; NUMBER OF SEQUENCES: 22

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Ciba-Geigy Corporation

; STREET: 7 Skyline Drive

; CITY: Hawthorne

; STATE: NY

; COUNTRY: USA

; ZIP: 10532

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/258,261B

; FILING DATE: 08-JUN-1994

; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/457,205

; FILING DATE: 01-JUN-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Elmer, James Scott

; REGISTRATION NUMBER: 36,129

; REFERENCE/DOCKET NUMBER: CGC 1506/CIP3

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 919-541-8614

; TELEFAX: 919-541-8689

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 28958 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHEetical: NO

; ANTI-SENSE: NO

US-08-258-261B-6

Alignment Scores:

Pred. No.: 0.574

Score: 126.50

Percent Similarity: 34.35%

Best Local Similarity: 22.35%

Query Match: 4.90%

DB: 1

Length: 28958

Matches: 95

Conservative: 51

Mismatches: 185

Indels: 114

Gaps: 18

US-09-830-144-4 (1-504) x US-08-258-261B-6 (1-28958)

QY 84 LeuSerAlaGluLeuLeuGlyGlnLeuAsnAlaGluHisAlaGluAlaAspValArg 103
DB 5506 CTGAAGCCCGCTCTCTCTCGGACACCTCCATTGGGACCTGTCGGCCGACCTCCG 5565
QY 104 ArgValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGlnSerIleAspAsp 123
DB 5566 GCGGTCCTT-----TCTCTCCAGGAC 5586
QY 124 Ala-----LeuAlaGluLysAlaSerLeuGlnSerGlnLeuProGluGlyValPro 140
DB 5587 GGCTGACCCCTGTCGCCCGCCCGGAAAGCTCATGCAAGCCCTCCACAAAGCCGCGCC 5646
QY 141 GlnHisGlnLeuProGlnTyrGlnLysIleLeuGluArgLeuLysThrLeuGluArg 160
DB 5647 ATGTCACCCCTCCGAGCTCCGAGGAGAAATCCGACCTTCACAGCCCTACGAAGGC 5706
QY 161 GluIleSerGlyAlaMetAlaValAlaValLeuLeuAsnAsnLysLeuTyrVal 180
DB 5707 CGAGCTAGCTGCGCGCC-----CTCATGAGGCTCTCTCCACCC 5745
QY 181 AlaAsnValGlyThrAsnArgAlaLeuLeu-----CysLysSerThrValAspGly 197
DB 5746 GTGCTGCGCTGGGATGAAGACCGGCTGTGGATGATCGCCCGCAGGCGAAAGCCCTCGA 5805
QY 198 LeuGlnValThrGlnLeuAsnValAspHisThrThrGluAsnGlu----- 212
DB 5806 CGAAAGACCAACGCGCTGCGCTGAGCCACGCTTCATCCCGCACATGAGCGGAATG 5865
QY 213 ---AspGluLeuPheArgLeuSerGln---LeuGlyLeuAspAlaGlyLysIleLysGln 230
DB 5866 CTCGACGACTTCGCGCGCTGCGCCAGAGCCCTACCTACCTCCGACGATCCCATC 5925
QY 221 ValGlyIleIleCysGlyGlnGlnSerThrArgArgIleGlyAspTyrLysValLysTyr 250
DB 5926 ATCTCCACAGTACCGCGCGCGCGCGCGCCAG-----GACCACGAG----- 5964
QY 251 GlyTyrThrAspIleAspLeuLeuSerAlaAlaLysSerLysProIleIleAlaGluPro 270
DB 5965 -----CTGCGCTCGGCC 5976
QY 271 GluIle-----HisGlyAlaGlnProLeuAspGlyValThrGlyPhe 284
DB 5977 GACTACTGGGCTCGCCACGTTCCGACACCGTCCGCTTCCTCGACGCGCTA----- 6027
QY 285 LeuValLeuMetSerGluGlyLeuTyrLysAlaLeuGluAlaHisGlyProGlyGln 304
DB 6028 -----CGTGCCCTTCACCGCGAAAGG--- 6048
QY 305 AlaAsnGlnGluIleAlaMetIleAspThrGluPheAlaLysGlnThrSerLeuAsp 324
DB 6049 -----GCAAGTGCTTTCTCGAGCTCGGGGCTCACGCTGCTCTCC 6090
QY 325 AlaValAlaGlnAlaValVal-AspArgValLys---ArgIleHisSerAspThrPheAl 343
DB 6091 GCCCTTGCGAAGACGCTCGGACAGAGCAAGGACGTCGCAATGCGCTCTCTCC 6150
QY 343 aserGlyGlyArgAla-----ArgPheCysProArgHisGluAspMetTh 359
DB 6151 ACCCTCCGCAAGGACGCGAGACGCGCGAGGCGTTACCGCGCGCC-----TCGGC 6201
QY 359 rLeuLeuValArgAsnPheGlyTyrProLeuGlyGluMetSerGlnPro----- 375
DB 6202 GCTCTCCACTCGCGAGCATCACCCGACTGAGCGCTTCTTCGCGCCCTTCGCTCCA 6261
QY 376 -----ThrProSerPro-----AlaProAlaAlaGlyArgValTyrPhe 389
DB 6262 CGCAAGGCTCTCTCCCACTATGCTTCGAGCGCGCGCTTCTTGCGCCGACGCTCTCC 6321
QY 389 oValSerValProTyrSerSerAlaGlnSerThrSerLysThrSerValThrLeuSerLe 409
DB 6322 A-----AGGCACCGCGCGCGCGAGCTCAGGCACTTGCTCCG 6357

QY 409 uValMetProSerGlnGlyGlnMetValAsnGlyAlaHisSerAlaSerThrLeuAspG 429
DB 6358 CTCGAGGGGGGCTCTGTGGCAAGCCATGACGCGGGGACCTCGATGGCTCAGCGGTCA 6417
QY 429 uAlaThrProThrLeuThrAsnGlnSerProThrLeuThrLeuGlnSerThrAsnThrH 449
DB 6418 CTCACGTCGACGCGGACGAGCGCGCGCGCTGCTCTCTCCACCTCTCG 6477
QY 449 sThrGlnSerSerSerSerSerSerAspGlyGlyLeuPheArgSerArgProAlaHis 469
DB 6478 AGCTTTCGCCAGAGCGCGCAAGACAGAGCA-----CGGTGACGCTTGCGCGCTAC 6528
QY 469 rLeuProProGly 473
DB 6529 CGTATCACTTGA 6541

Search completed: December 10, 2002, 02:07:20
Job time : 115.518 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 9, 2002, 23:04:29 ; Search time 72.5983 Seconds
(without alignments)
2707.033 Million cell updates/sec

Title: US-09-830-144-4
Perfect score: 2580
Sequence: 1 MAARRSILDSQQPSWTTD.....AEFFRLMSVDHGQSVTAP 504

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 350425 segs, 19496369 residues

Total number of hits satisfying chosen parameters: 700850

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+.p2n.model -DEV=xip
-Q=/cgn2_1/USPRO.spool/US09830144/runat_04122002.141354.2307/app_query.fasta_1.1422
-DB=Published.Applications.NA -OPW=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 MATRIX=blsune62
-TRANS=human40.cdd -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HRAPIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09830144@cgn2_1.1.51@runat_04122002.141354.2307
-NCPU=6 -ICPU=3 -NO XLPXY -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published.Applications.NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2580	100.0	1560	9	US-10-158-895-1
2	2580	100.0	1560	12	US-10-123-427-1
3	2580	100.0	1569	9	US-10-158-895-10
4	2577	99.9	1568	9	US-10-158-895-42

	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	
	2575	543.5	409	365	323.5	271.5	261	196.5	190.5	188.5	186	178	175	171	171	171	171	171	171	171	171	171	171	169.5	168.5	167.5	164.5	164.5	158	156.5	156.5	155	155	154.5	154.5	151.5	146	145.5	142	139.5	138.5	
	99.8	21.1	16877	14.1	12.5	10.5	10.1	7.6	7.4	7.3	7.2	6.9	6.8	6.6	6.6	6.6	6.6	6.6	6.6	6.6	6.6	6.6	6.6	6.6	6.5	6.5	6.4	6.4	6.1	6.1	6.0	6.0	6.0	6.0	6.0	5.9	5.7	5.6	5.5	5.4	5.4	
	1560	1560	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
	US-10-123-427-5	US-09-764-877-3349	US-09-925-300-330	US-09-864-761-17553	US-09-864-761-1769	US-09-864-761-1770	US-09-864-761-1754	US-09-938-842A-2158	US-09-938-842A-2158	US-09-938-842A-210	US-09-770-445-6	US-09-935-124A-1	US-09-938-842A-1634	US-09-973-941-3	US-09-973-963-3	US-09-973-964-3	US-09-973-964-3	US-09-973-964-3	US-09-973-964-3	US-09-973-964-3	US-09-973-964-3	US-09-973-964-3	US-09-973-964-3	US-09-973-964-3	US-09-938-842A-1547	US-09-925-300-715	US-09-938-842A-1434	US-09-860-351-3	US-09-860-351-3	US-09-938-842A-736	US-09-828-302-9	US-09-880-107-1608	US-09-920-300A-1714	US-10-033-528-1714	US-09-938-842A-212	US-09-938-842A-1595	US-09-938-842A-1208	US-09-938-842A-1818	US-09-828-302-10	US-09-938-842A-1925	US-09-938-842A-1133	US-09-938-842A-1948
	Sequence 5, Appli	Sequence 3349, Ap	Sequence 330, App	Sequence 17553, A	Sequence 769, App	Sequence 770, App	Sequence 1754, A	Sequence 2158, Ap	Sequence 958, App	Sequence 210, App	Sequence 6, Appli	Sequence 1, Appli	Sequence 1634, Ap	Sequence 3, Appli	Sequence 3, Appli	Sequence 3, Appli	Sequence 3, Appli	Sequence 3, Appli	Sequence 3, Appli	Sequence 3, Appli	Sequence 3, Appli	Sequence 3, Appli	Sequence 3, Appli	Sequence 3, Appli	Sequence 1547, Ap	Sequence 715, App	Sequence 1434, Ap	Sequence 1, Appli	Sequence 736, App	Sequence 9, Appli	Sequence 1608, Ap	Sequence 1714, Ap	Sequence 1114, Ap	Sequence 212, App	Sequence 1595, App	Sequence 1818, Ap	Sequence 1818, Ap	Sequence 10, Appl	Sequence 1925, Ap	Sequence 1133, Ap	Sequence 1948, Ap	

ALIGNMENTS

RESULT 1
US-10-158-895-1
Sequence 1, Application US/10158895
Patent No. US20020155624A1
GENERAL INFORMATION:
APPLICANT: ONO, KOICHIRO
APPLICANT: TSUCHIYA, MASAYUKI
APPLICANT: TOSHIHIKO
TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
FILE REFERENCE: 053466/0278
CURRENT APPLICATION NUMBER: US/10/158, 895
CURRENT FILING DATE: 2002-06-03
PRIOR APPLICATION NUMBER: US/09/529, 279
PRIOR FILING DATE: 2000-04-11
PRIOR APPLICATION NUMBER: PCT/JP98/04796
PRIOR FILING DATE: 1998-10-22
PRIOR APPLICATION NUMBER: JP 9/290188
PRIOR FILING DATE: 1997-10-22
NUMBER OF SEQ ID NOS: 48
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 1560
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (30)..(1541)
US-10-158-895-1

Alignment Scores:

Pred. No.: 2,32e-229 Length: 1560
Score: 2580.00 Matches: 504
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-830-144-4 (1-504) x US-10-158-895-1 (1-1560)

```
Qy 1 MetAlaAlaGlnArgSerLeuLeuGlnSerGluGlnGlnProSerTrpThrAspAsp 20
Db 30 ATGGCGCGCAGAGGAGGAGCTTGTGTCAGAGTGAGCAGCAGCCAGCTGGAGCATGAC 89
Qy 21 LeuProLeuCysHisLeuSerGlyValGlySerAlaSerAsnArgSerTyrSerAlaAsp 40
Db 90 CTGCTCTCTGCGCACCTCTCTGGGGTGTGCTCAGCCTCCAAACCGCAGCTACTCTGCTGAT 149
Qy 41 GlyLysGlyThrGluSerHisProGluAspSerTrpLeuLysPheArgSerGluAsn 60
Db 150 GGCAAGGCACTCAGAGCCACCGCCAGAGGACAGCTGGCTCAAGTTCAGGAGTGAGAAC 209
Qy 61 AsnCysPheLeuTyrGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsnPheVal 80
Db 210 AACTGTCTCTGTATGGGGTCTTCAACGGCTATGATGGCAACCGAGTGACCAACTTCGTG 269
Qy 81 AlaGlnArgLeuSerAlaGluLeuLeuGlyGlnLeuAsnAlaGluHisAlaGluAla 100
Db 270 GCCACGGGCTGTCCGAGAGCTCTCTGCTGGCCACCTGAATGCCGAGCAGCCGAGGCC 329
Qy 101 AspValArgArgValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGluSer 120
Db 330 GATGTGCGGCGTGTGCTGTCAGGCGCTTCGATGTGCTGAGAGGAGCTTCTCTGAGTCC 389
Qy 121 IleAspAspAlaLeuAlaGluLysAlaSerLeuGlnSerGlnLeuProGluGlyValPro 140
Db 390 ATTGACAGCCCTTGGCTGAGAGGCAAGCCTCCAGTCGCAATGTCAGAGGAGTCCCT 449
Qy 141 GlnHisGlnLeuProGlnTyrGlnLysIleLeuGluArgLeuLysThrLeuGluArg 160
Db 450 CAGCACCGCTGCTCTCAGTATCAGAGATCTTTGAGAGACTCAAGACGTTAGAGAG 509
Qy 161 GluIleSerGlyGlyAlaMetAlaValAlaValLeuLeuAsnAsnLysLeuTyrVal 180
Db 510 GAAATTTCCGGAGGGCCATGCGCGTGTGGCGGTCTTCTCAACAACCAAGCTCTACGTC 569
Qy 181 AlaAsnValGlyThrAsnArgAlaLeuLeuCysLysSerThrValAspGlyLeuGlnVal 200
Db 570 GCCAATGTCGGTACAAACCGTGCACCTTTATGCAATCGACAGTGGTGGTGGAGTG 629
Qy 201 ThrGlnLeuAsnValAspHisThrThrGluAsnGluAspGluLeuPheArgLeuSerGln 220
Db 630 ACACAGCTGAACGTGGACCAACACACAGAGAACAGAGATGAGCTCTCCGTCTTTCCGAG 689
Qy 221 LeuGlyLeuAspAlaGlyLysIleLysGlnValGlyIleLysCysGlyGlnGluSerThr 240
Db 690 CTGGCTTGGATGCTGGAAGATCAAGCAGGTGGGGATCATCTGTGGCAGGAGAGCACC 749
Qy 241 ArgArgIleGlyAspTyrLysValLysTyrGlyTyrThrAspIleAspLeuLeuSerAla 260
Db 750 CGCGCGATCGGGGATTACAAGTTAAATATGGCTACACGACATTGACCTTCTCAGCGCT 809
Qy 261 AlalysSerLysProIleIleAlaGluProGluIleHisGlyAlaGlnProLeuAspGly 280
Db 810 GCCAAGTCCAAACCAATCATCGACAGCCAGAAATCCATGGGCACACCGCTGGATGGG 869
Qy 281 ValThrGlyPheLeuValLeuMetSerGluGlyLeuTyrLysIleAlaLeuAlaHis 300
Db 870 GTGACGGGCTCTTGTGTGCTGATGTCGAGGGGTGTGACAAAGCCCTAGAGGAGGCCCAT 929
Qy 301 GlyProGlyGlnAlaAsnGlnGluIleAlaAlaMetIleAspThrGluPheAlaLysGln 320
```

```
Db 930 GGGCTTGGGCGAGCCCAACAGGAGATTGTGCGATGATTGACACTGATTGTTGCCAAGCAG 989
Qy 321 ThrSerLeuAspAlaValAlaGlnAlaValValAspArgValLysArgIleHisSerAsp 340
Db 990 ACCTCCCTGGACAGTGGCCCGCTGCTGAGCCGCGTGAAGCGCATCCACACGCGAC 1049
Qy 341 ThrPheAlaSerGlyGlyGluArgAlaArgPheCysProArgHisGluAspMetThrLeu 360
Db 1050 ACCTTCGCCAGTGGTGGGAGCGTGCAGGTTCTGCCCCCGGACCGAGGACATGACCTG 1109
Qy 361 LeuValArgAsnPheGlyTyrProLeuGlyGluMetSerGlnProThrProSerProAla 380
Db 1110 CTAGTGAGGAACCTTGGCTACCCGCTGGGCGAAATAGCCAGCCACACCGAGCCGAGCC 1169
Qy 381 ProAlaAlaGlyArgValTyrProValSerValProTyrSerSerAlaGlnSerThr 400
Db 1170 CCAGCTCAGAGGAGCAGAGTGATCCCTGTGTGTGCCATCTCCAGCGCCCGAGACACC 1229
Qy 401 SerLysThrSerValThrLeuSerLeuValMetProSerGlnGlyGlnMetValAsnGly 420
Db 1230 AGCAAGACACGAGTGACCTCTCCCTGTGTCATGCTCCAGGGCCAGATGGTCAACGGG 1289
Qy 421 AlaHisSerAlaSerThrLeuAspGluAlaThrProThrLeuThrAsnGlnSerProThr 440
Db 1290 GCTCACAGTGTCTCCACCTGGAGAGCCACCCCAACCTCCACCAACCAAGCCGAGCC 1349
Qy 441 LeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSerSerSerAspGly 460
Db 1350 TTAACTCTGTCAGTCCACCAACAGCAGCAGCAGCAGCAGCTCCAGCTCTGACGAGGC 1409
Qy 461 LeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGlyArgValGluPro 480
Db 1410 CTCCTTCGCTCCCGCCCGCCCACTGCTCCGCTCCGCGAGGAGCGTCTGTTGAGCCC 1469
Qy 481 TyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGlyGluGlnSerVal 500
Db 1470 TATGTGAGCTTTGCTGAGTTTACCGCTCTGGAGCGTGGACCATGCGAGCAGAGCGGTG 1529
Qy 501 ValThrAlaPro 504
Db 1530 GTGACAGCACCG 1541

RESULT 2
US-10-123-427-1
; Sequence 1, Application US/10123427
; Patent No. US20020119525A1
; GENERAL INFORMATION:
; APPLICANT: MATSUMOTO, Kunihiro
; NISHIDA, Eisuke
; TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/123,427
; FILING DATE: 17-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/406,854
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US/08/752,891
; FILING DATE: 20-NOV-1996
; APPLICATION NUMBER: JP 8-300856
```

FILING DATE: 28-OCT-1996
APPLICATION NUMBER: JP 8-126282
FILING DATE: 24-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 17981/111
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1560 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 30..1541
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 30..1541
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-123-427-1
Alignment Scores:
Pred. No.: 2,32e-229 Length: 1560
Score: 2580.00 Matches: 504
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
US-09-830-144-4 (1-504) x US-10-123-427-1 (1-1560)
QY 1 MetaAlaGlnArgSerLeuGlnSerGlnGlnProSerTrpThrAspAsp 20
Db 30 ATGGCGGGCGAGAGAGAGAGCTTCTGTCAGAGTACGACGACGACCACTGGACAGATGAC 89
QY 21 LeuProLeuGlyHisLeuSerGlyValGlySerAlaSerAsnArgSerTrpSerAlaAsp 40
Db 90 CTGCTCTCTCCACCTCTCTGGGTTGGCTCAGCTCCAACTGCACTACTGCTGAT 149
QY 41 GlyGlyGlyThrGlnSerHisLeuProGlnAspSerTrpLeuGlyPheAlaGlySerGlnAsn 60
Db 150 GGCAAGGGGCACTGAGAGCCGCCGAGAGGACGCTGCTCAAGTTCAGAGAGGAGAAC 209
QY 61 AsnGlyPheLeuTrpGlyValPheAsnGlyTrpAspGlyAsnArgValThrAsnPheVal 80
Db 210 AACTGCTTCTGTAAGGGGTTCTTCAACGGCTATGATGGCAACGAGTACCACTTCTGTG 269
QY 81 AlaGlnArgLeuSerAlaGlnLeuLeuGlnGlnLeuAsnAlaGlnHisAlaGlnAla 100
Db 270 GCCACGCGGCTGTCCGCGAGAGCTCTGCTGGCCAGCTGATGCGAGACGCGGAGGCC 329
QY 101 AspValArgArgValLeuLeuGlnAlaPheAspValValGlnArgSerPheLeuGlnSer 120
Db 330 GATGCGCGCTGTCTGCTGCTGCTGAGGCTTCGATGTGTGGAGAGAGCTTCTGAGATCC 389
QY 121 IleAspAspAlaLeuAlaGlnValAlaSerLeuGlnSerGlnLeuProGlnGlyValPro 140
Db 390 ATTGACGAGGCTTGTGGTGAAGGCAAGCTTCCAGCTGCAATTGCGCAGAGGAGTCCCT 449
QY 141 GlnHisGlnLeuProGlnGlnTrpGlnHisIleLeuGlnArgLeuGlyThrLeuGlnArg 160
Db 450 CAGCACGAGCTTCCCTCTCAAGATCAGAAAGATCTTTGAGAGACTCAAGACTTTAGAGAG 509
QY 161 GlnIleSerGlyGlyAlaMetAlaValAlaValAlaValLeuLeuAsnAsnLeuTrpVal 180
Db 510 GAAATTTGGGAGGGGCGATGCGCTTGTGGGCGTCTTCTCAACAAAGCTCTACGTC 569
QY 181 AlaAsnValGlyThrAsnArgAlaLeuLeuGlySerTrpValAspGlyLeuGlnVal 200

Db 570 GCCATGTGCGTACAAACCGGCACTTTATGCAAAATGCAAGTGGTGTCCAGTGG 629
QY 201 ThrGlnLeuAsnValAspHisLeuThrGlnAsnGlnAspGlnLeuPheArgLeuSerGln 220
Db 630 ACACAGCTGAACGTGGACCAACACACAGAGAACAGATGATGCTTCTCCGCTTCCGAC 689
QY 221 LeuGlyLeuAspAlaGlyValIleValSerGlnValGlyIleIleIleIleIleIleIleIle 240
Db 690 CTGGCTTGGATGCTGGAAAGATCAGACAGTGGGATCATCTGTGGGACAGAGACACC 749
QY 241 ArgArgIleGlyAspTrpValValIleValIleValIleValIleValIleValIleValIle 260
Db 750 CGGCGGATCGGGGATTCACAAAGTTAAATATGCTTACACGGGATTTGACCTTCTCAGCGCT 809
QY 261 AlaySerLeuProIleIleAlaGlnProGlnIleHisGlyValAlaGlnProLeuAspGly 280
Db 810 GCCAAGTCCAAACCAATATGACAGAGCCAGAAATCCATGGGACACGCCCTGGATGG 869
QY 281 ValThrGlyPheLeuValLeuMetSerGlnGlyLeuTrpValAlaLeuGlnAlaAlaHis 300
Db 870 GTGACGGGCTTCTGTCTGATGTGAGGGGTTGTACAAAGCCCTTAGAGGACGCCAT 929
QY 301 GlyProGlyGlnAlaAsnGlnGlnIleAlaAlaMetIleAspThrGlnPheAlaGlyGln 320
Db 930 GGGCTGGGCGAGGCCAACAGAGAGATGCTGCTGATGTGACACTGATTTGCCAAGCAG 989
QY 321 ThrSerLeuAspAlaValAlaGlnAlaValAlaAspArgValIleHisSerAsp 340
Db 990 ACCCTCTTGGAGAGAGTGGCCCGGCTCGTGACCGGGTGAAGCGATCACAGGCAC 1049
QY 341 ThrPheAlaSerGlyGlyValArgAlaArgPheCysProArgHisGlyAspMetThrLeu 360
Db 1050 ACCCTGCGCAGTGTGGGAGGCTCCAGCTTCTCCCGCGGACGAGACATACCCCTG 1109
QY 361 LeuValArgAsnPheGlyTrpProLeuGlyGlnMetSerGlnProThrProSerProAla 380
Db 1110 CTAGTAGAGAACTTTGGCTACCCGCTGGCGAAATGAGCCAGCCACAGCCAGCCACCC 1169
QY 381 ProAlaAlaGlyValArgValTrpProValSerValProTrpSerSerAlaGlnSerThr 400
Db 1170 CCAGCTGACGAGGAGAGAGTGAACCTGTGTGTGTGTCATCTCCAGGCGCCAGAGGCC 1229
QY 401 SerLeuThrSerValThrLeuSerLeuValMetProSerGlnGlyMetValAsnGly 420
Db 1230 AGCAGACACGCTGACCTTCTCTCTGTCATGCTCTCCAGGCGCAATGTCACCGG 1289
QY 421 AlaHisSerAlaSerThrLeuAspGlnAlaThrProThrLeuThrAsnGlnSerProThr 440
Db 1290 GCTCAGAGTCTTCCACCTGAGAGAGCAACCCACCTCACCACCAACCAAGCCGAC 1349
QY 441 LeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSerSerAlaGly 460
Db 1350 TTAACTCTCAGTCCACCAACAGCAGACGCGAGACAGCTCCAGCTGAGCGAGAGC 1409
QY 461 LeuPheArgSerArgProAlaHisSerLeuProProGlyGlnAspGlyArgValGlnPro 480
Db 1410 CTCTTCCGCTCCCGGCGCCCACTGCTCCGCTGGGAGAGAGGTCGTATTAGCCC 1469
QY 481 TrpValAspPheAlaGlnPheTrpArgLeuTrpSerValAspHisGlyGlnGlnSerVal 500
Db 1470 TATGTGACTTTGTGATGTTTACCGCTCTGGAGCGTGAACATGGGAGAGCGCGT 1529
QY 501 ValThrAlaPro 504
Db 1530 GTGACAGCACCG 1541
RESULT 3
US-10-158-895-10
; Sequence 10, Application US/10158895
; Patent No. US20020155624A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO

; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 1569
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)..(1557)
US-10-158-895-10
Alignment Scores:
Pred. No.: 2,34e-229 Length: 1569
Score: 2580.00 Matches: 504
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-09-830-144-4 (1-504) x US-10-158-895-10 (1-1569)
QY 1 MetAlaGlnArgSerLeuLeuGlnSerGluGlnGlnProSerTrpThrAspAsp 20
Db 7 ATGGCGCGCAGAGGAGGCTTGTGCGAGGTGAGCAGCAGCAAGCTCGACAGATGAC 66
QY 21 LeuProLeuGlnSerGlnSerGlnValGlySerAlaSerAsnArgSerTrpSerAlaAsp 40
Db 67 CTGCTCTCTGCCACCTCTCTGGGGTGGCTCAGCCCTCAACCGCAGCTACTCTGCTGAT 126
QY 41 GlyLysGlyThrGluSerHisProProGluAspSerTrpLeuLysPheArgSerGluAsn 60
Db 127 GGCAAGGGCACTCAGAGCCACCGCCAGAGGACAGCTGGCTCAAGTTCAAGAGTGAAC 186
QY 61 AsnCysPheLeuTyrGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsnPheVal 80
Db 187 AACTGTCTCTGTATGGGTCTTCAACGGCTATGATGGCAACCGAGTGACCAACTTCGTG 246
QY 81 AlaGlnArgLeuSerAlaGluLeuLeuLeuGlyGlnLeuAsnAlaGluHisAlaGluAla 100
Db 247 GCCACGGCTGTCCGACAGCTCTCTGTGGCCAGCTGAATGCCGAGCACCAGGAGGCC 306
QY 101 AspValArgValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGluSer 120
Db 307 GATGTGGCGCTGTGTGTGCGAGGCTTCGATGTGTGGAGGAGCTTCTCTGGAGTCC 366
QY 121 IleAspAlaLeuAlaGluLysAlaSerLeuGlnSerGlnLeuProGluGlyValPro 140
Db 367 ATTGACGAGCGCTTGGCTGAGAGGCAAGCCCTCCAGTCGCAATTTGCCAGGGAGTCCCT 426
QY 141 GlnHisGlnLeuProProGlnTyrGlnLysIleLeuGluArgLeuLysThrLeuGluArg 160
Db 427 CAGCACAGCTGCCTCTCTCAGTATCAGAAGATCTTTGAGAGACTCAAGACGTTAGAGG 486
QY 161 GluIleSerGlyGlyAlaMetAlaValAlaValAlaLeuLeuAsnLysLeuTyrVal 180
Db 487 GAAATTCGGAGGGCCATGGCGGTGTGGCGGTCTTCTCAACAACAAGCTCTACGTC 546
QY 181 AlaAsnValGlyThrAsnArgAlaLeuLeuCysLysSerThrValAspGlyLeuGlnVal 200
Db 547 GCCAATGTCGGTACAAACCGTCACCTTTTATGCAATTCAGACAGTGGTGGTGGAGGTG 606

QY 201 ThrGlnLeuAsnValAspHisThrGluAsnGluAspGluLeuPheArgLeuSerGln 220
Db 607 ACACAGCTGAACGTTGGACACACACAGAGAACGAGGATGAGCTTCCGCTTTTCGACG 666
QY 221 LeuGlyLeuAspAlaGlyLysIleValGlyIleLeuCysGlyGlnGluSerThr 240
Db 667 CTGGCTTGGATGCTGGAAAGATCAACAGCAGGTGGGATCATCTGTGGCAGGAGGACCC 726
QY 241 ArgArgIleGlyAspTyrLysValLysValTyrGlyTyrThrAspIleAspLeuSerAla 260
Db 727 CGGCGATCGGGATTAACAAGTTAAATATGGCTACACGACATTGACCTTCTCAGCGCT 786
QY 261 AlaLysSerLysProIleIleAlaGluProGluIleHisGlyAlaGlnProLeuAspGly 280
Db 787 GCCAAGTCCAAACCAATCATCGCAGAGCCAGAAATCCATGGGCACAGCCGCTGGATGGG 846
QY 281 ValThrGlyPheLeuValLeuMetSerGluGlyLeuTyrLysAlaLeuGluAlaHis 300
Db 847 GTGACGGCTTCTTGTGCTGATGTCGAGGGGTGTGTACAGGCCCTTAGAGGACGCCAT 906
QY 301 GlyProGlyGlnAlaAsnGlnGluIleAlaMetIleAspThrGluPheAlaLysGln 320
Db 907 GGGCTTGGCAGGCCAACAGGAGATTGCTCGATGATTGACACTGAGTTTGCCAGCAG 966
QY 321 ThrSerLeuAspAlaValAlaGlnAlaValAspArgValLysArgIleHisSerAsp 340
Db 967 ACCTCCCTTGGACGCGAGTGGCCAGGCCGTCGTGGACCGGTGAAGCGCATCCACAGCGAC 1026
QY 341 ThrPheAlaSerGlyGlyGluArgAlaArgPheCysProArgHisGluAspMetThrLeu 360
Db 1027 ACCTTCGCCAGTGTGGGGAGCGTCCAGGTTCTGCCCCCGGCACGAGGACATGACCCCTG 1086
QY 361 LeuValArgAsnPheGlyTyrProLeuGlyGluMetSerGlnProThrProSerProAla 380
Db 1087 CTAGTAGGAACTTTTGGCTACCCGCTGGCGAAATGAGCAGGCCACACCGAGCCAGGCC 1146
QY 381 ProAlaAlaGlyArgValTyrProValSerValProTyrSerSerAlaGlnSerThr 400
Db 1147 CCAGCTGAGAGGAGCAGGTGATCCCTGTGTGTGTCATCTCCAGCGCCAGAGCACC 1206
QY 401 SerLysThrSerValThrLeuSerLeuValMetProSerGlnGlyGlnMetValAsnGly 420
Db 1207 AGCAAGACCGAGCGTACCCCTCTCCCTTGTCTATGCCCTCCCGAGGCCAGATGGTCAACGG 1266
QY 421 AlaHisSerAlaSerThrLeuAspGluAlaThrProThrLeuThrAsnGlnSerProThr 440
Db 1267 GCTCACAGTGTCTCCACCTGGACGAAGCACCCCAACCTCCACCAACCAAGCCGAGCC 1326
QY 441 LeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSerSerSerSerGlyGly 460
Db 1327 TTAACCTTGAGTCCCAACACAGCACACGAGCAGCAGCTCCAGCTCTGACGAGGC 1386
QY 461 LeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGlyArgValGluPro 480
Db 1387 CTCTTCGCTCCCGCCCGCCACTCGCTCCCGCTTGGCGAGGACCGTCTGTGTGAGCCC 1446
QY 481 TyrValAspPheAlaGluPheTyrArgLeuTyrSerValAspHisGlyGluGlnSerVal 500
Db 1447 TATGTGACTTTGTGAGTTTATCCGCTCTGAGCGGTGGACCATGGCAGCAGAGCGGTG 1506
QY 501 ValThrAlaPro 504
Db 1507 GTGACAGCACCG 1518

RESULT 4

US-10-158-895-42
; Sequence 42, Application US/10158895
; Patent No. US2002015624A1
; GENERAL INFORMATION:
; APPLICANT: OHTOMO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES

FILE REFERENCE: 053466/0278
 CURRENT APPLICATION NUMBER: US/10/158,895
 CURRENT FILING DATE: 2002-06-03
 PRIOR APPLICATION NUMBER: US/09/529,279
 PRIOR FILING DATE: 2000-04-11
 PRIOR APPLICATION NUMBER: PCT/JP98/04796
 PRIOR FILING DATE: 1998-10-22
 PRIOR APPLICATION NUMBER: JP 9/290188
 PRIOR FILING DATE: 1997-10-22
 NUMBER OF SEQ ID NOS: 48
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 42
 LENGTH: 1568
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (11)..(1549)
 US-10-158-895-42

Alignment Scores:

Pred. No.: 4,43e-229 Length: 1568
 Score: 2577.00 Matches: 503
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.80% Mismatches: 0
 Query Match: 99.88% Indels: 0
 Gaps: 0

US-09-830-144-4 (1-504) x US-10-158-895-42 (1-1568)

Qy 1 MetIaIaIaGlnArgSerLeuLeuGlnSerGlnGlnProSerThrThrAsp 20
 Db 38 ATGGGGGCGAG 97
 Qy 21 LeuProLeuCyHisLeuSerGlyValGlySerAlaSerAsnArgSerTyrSerAla 40
 Db 98 CTGCTCTCTGCGACCTCTCTGGGGGTGGCTCAGCTCCACCGAGAGACTCTGAT 157
 Qy 41 GlyIysGlyThrGlnSerHisProProGluAspSerThrLeuLeuPheArgSerGlu 60
 Db 158 GCGAAGGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 217
 Qy 61 AsnCySPheLeuTyrGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsn 80
 Db 218 AACTGCTTCTGTATGGGGGTCTTCAAGGCTATGATGAGCAACGAGTGCAT 277
 Qy 81 AlaGlnArgLeuSerAlaGlnLeuLeuGlyGlnLeuAsnAlaGlnHisAlaGln 100
 Db 278 GCCCAGCGGCTGTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 337
 Qy 101 AspValArgValAlaLeuGlnAlaPheAspValValGlnArgSerPheLeuGln 120
 Db 338 GATGGCGGCGCTGTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 397
 Qy 121 IleAspAspAlaLeuAlaGlnValAlaSerLeuGlnSerGlnLeuProGlnVal 140
 Db 398 ATTGACGAGCGCTTGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 457
 Qy 141 GlnHisGlnLeuProProGlnTyrGlnValIleLeuGlnArgLeuTyrThrLeu 160
 Db 458 CAGCAGCAGAGCTCTCTCAGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 517
 Qy 161 GluIleSerGlyGlyAlaMetAlaValAlaValLeuLeuAsnValLeuTyrVal 180
 Db 518 GAAATTTTGGGAGGGGCGATGGCGGTGGCGGCTCTTCAACCAAGCTTCAAGC 577
 Qy 181 AlaAsnValGlyTyrAsnArgAlaLeuLeuCysLeuSerThrValAspGlyLeuGln 200
 Db 578 GCCATGTGCGTACCAACCGGCACTTTTATGCAATCCAGAGTGGATGGTGGAG 637
 Qy 201 ThrGlnLeuAsnValAspHisThrThrGluAsnGluAspGluLeuPheArgLeuSer 220
 Db 638 ACACAGCTGAACGTGACCAACACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 697

Qy 221 LeuGlyLeuAspAlaGlyIysIleLeuGlnValGlyIleCysGlyGlnLeuSerThr 240
 Db 698 CTGGCTTGGAGTGTGGAGAGAGATCAAGCGAGTGGGATATCTTGGGAGAGAGAG 757
 Qy 241 ArgArgIleGlyAspTyrIysValIysTyrGlyTyrThrAspIleAspLeuSerAla 260
 Db 758 CCGCGGATCGGGGATTCAGAGGTTAAATATGGCTACACGAGCATTCCTCTCAGCC 817
 Qy 261 AlAspSerIysProIleIleAlaGlnProGlnIleHisGlyAlaGlnProLeuAsp 280
 Db 818 GCAAGTCCAAATCAATCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 877
 Qy 281 ValThrGlyPheLeuValLeuMetSerGlnGlyLeuTyrValAlaLeuGlnAlaHis 300
 Db 878 GTGAGGGCTTCTTGGTGTGATGTGAGAGGGGTGTACCAAGCCCTAGAGAGCCAT 937
 Qy 301 GlyProGlyGlnAlaAsnGlnIleAlaIleAlaMetIleAspThrGluPheAlaGln 320
 Db 938 GGGCTGGGAGGCGCAACAGAGATTGCTGCATGATTTGACATGAGTTGGCCAAGC 997
 Qy 321 ThrSerLeuAspAlaValAlaGlnAlaValAlaAspArgValIysArgIleHisSer 340
 Db 998 ACTCCCTGAGCGAGTGGCGCCAGGCGTGTGGAGCCGGGTGAAGCGCATCCAGCG 1057
 Qy 341 ThrPheAlaSerGlyGlyGlnValAlaArgPheCysProArgHisGluAspMetThr 360
 Db 1058 ACCTTGGCAGTGTGGGAGCGTGGCAGGTTCTGCCCGGAGAGAGAGAGAGAG 1117
 Qy 361 LeuValArgAsnPheGlyTyrProLeuGlyGlnMetSerGlnProThrProSerPro 380
 Db 1118 CTAGTGAAGAACTTGGCTACCCGCTGGCCAAATGAGCAGCCAGCCAGCCAGCC 1177
 Qy 381 ProAlaAlaGlyIysValArgValTyrProValSerValProTyrSerSerAlaGln 400
 Db 1178 CAGCTGAG 1237
 Qy 401 SerIysThrSerValThrLeuSerLeuValMetProSerGlnGlyIleMetValAsn 420
 Db 1238 AGCAAGACCAAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1297
 Qy 421 AlaHisSerAlaSerThrLeuAspGlnAlaThrProThrLeuThrAsnGlnSerPro 440
 Db 1298 GCTCAGATGCTTCCACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1357
 Qy 441 LeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSerAspGly 460
 Db 1358 TTAACTTGCAGTCCACCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1417
 Qy 461 LeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGlyArgValGln 480
 Db 1418 CTCTCCGCTCCGAGCGGCGCCACTCGCTCCGCTGGAGAGAGAGAGAGAGAG 1477
 Qy 481 TyrValAspPheAlaGlnPheTyrArgLeuTyrSerValAspHisGlnGlnSerVal 500
 Db 1478 TATGAGACTTGTGCTGATTTTACCGCTCTGAGAGGTGACCATGAGAGAGAG 1537
 Qy 501 ValThrAlaPro 504
 Db 1538 GTGACAGCACCG 1549

RESULT 5
 US-10-123-427-5
 Sequence 5, Application US/10123427
 Patent No. US20020119525A1
 GENERAL INFORMATION:
 APPLICANT: MATSUMOTO, Kunihiko
 NISHIDA, Eisuke
 TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR
 NUMBER OF SEQUENCES: 8
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Foley & Lardner
 STREET: 3000 K Street, N.W., Suite 500

Qy 461 LeupheArSerArgProAlaHisSerLeuProGlyGluAspGlyArgValGluPro 480
Db 1410 CTCCTCCGCTCCCGCCCGCCCACTGCTCCGCTGGGAGAGACGATGTTGAGCCC 1469
Qy 461 TyrValAspPheAlaGluPheTyrArgLeuTyrSerValAspHisGlyGluGlnSerVal 500
Db 1470 TATGAGACTTGTGCTGAGTTTACCGCTCTGGAGCGTGCACCATGGCGAGAGACGCTG 1529
Qy 501 ValThrAlaPro 504
Db 1530 GTGACAGCACCG 1541

RESULT 6

US-09-764-877-3349
; Sequence 3349, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3349
; LENGTH: 16877
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-3349

Alignment Scores:

Prod. No.: 7,18e-40 Length: 16877
Score: 543.50 Matches: 295
Percent Similarity: 18.86% Conservative: 43
Best Local Similarity: 16.46% Mismatches: 107
Query Match: 21.07% Indels: 1348
Gaps: 19

US-09-830-144-4 (1-504) x US-09-764-877-3349 (1-16877)

Qy 12 GluGlnGlnProSerTyrThrAspAspLeuProLeuGlyHisLeuSerGlyValGlySer 31
Db 1 GAGCGAGCGCCAGCTGACGATGACCTCTCTGCGACCTCTGCGGTTGCTCA 60
Qy 32 AlaSerAsnArgSerTyrSerAlaAspGlyValGlyHisLeuSerProProGluAsp 51
Db 61 GCCTCAACCCGACGACTCTGCTGATGCAAGGCACTGAGAGCCACCCGCAAGAGAC 120
Qy 52 SerTyrLeuLysPhe----- 56
Db 121 AGCTGGCTCAAGTTCAGGTGTGTGTGCACGATTTCTGTTGTGGCCCGGAGATTGTT 180
Qy 56 ----- 56
Db 181 GGTTCGACAGAGAAAGACACGACCTTGACGCTTTCTGATGGGCTTGCCAGTGAC 240
Qy 56 ----- 56
Db 241 ATGTGGCCGTGAGAGGTGCTGCTGCTGCTTGTGCAAGGGCTGCTGATGGGT 300
Qy 56 ----- 56
Db 301 AGCGTAGCATGGGAGAGAGGGTATCCAGAAATGTCAATAGCCAGAGTGAATGATGGCTA 360
Qy 56 ----- 56
Db 361 AAGCAGGGGACCCAGAGAGGCCCTGTAAGCTGAGCTGCTGCTTAACTCCCA 420
Qy 56 ----- 56
Db 421 ATTCTTTCCTTTCTCTCTCTCACTCCGTGAGACCTGAGTCTCAGGCTCCCTCTGCCC 480

Qy 57 -----ArgSerGluAsnAsnGlyPheLeuTyrGlyValPheAsnGlyTyrAsp 72
Db 481 TCTCCCTTCCAGAGTGTGAACCACTGCTCTGTATGGGCTTTCACAGGCTATGAT 540
Qy 73 GlyAsnArgValThrAsnPheValAlaGlnArgLeuSerAlaGlyLeuLeuGlnGlyGln 92
Db 541 GGCACCCGAGTGAACCACTTGTGGCCAGCGGCTGTCCGAGAGCTCTGCTGGGCGAG 600
Qy 93 LeuAsnAlaGluHisAlaGluAlaAspValArgArgValLeuLeuGln----- 108
Db 601 CTGATGCGCAGACACCGCCGAGGCGGATGTGCGGCTGTCTCTGTGAGGTATGTTGCGG 660
Qy 108 ----- 108
Db 661 GGGCCACAGTGAACCCAGCACATCATGTCCCAACCCAGAGCTTGGCCCTGACCTC 720
Qy 108 ----- 108
Db 721 TAGCATGTTGCCAGGGTTGGTGTGAAGATCTTCGCGCCCTTACCCAGTAGAGAGC 780
Qy 108 ----- 108
Db 781 AGCTCCAGCTAGGCCCCCAACCAAGAGTCCAGAGACCAGCAGCAGCAGCAGCAGCAG 840
Qy 108 ----- 108
Db 841 CCTGAATGCTGGCAGAGAGGTTACTGCTGAGACGGGGGATTAGGAGTGGAGCTT 900
Qy 108 ----- 108
Db 901 GGAGAGAGTGTGAGTGTGGAGCAGGCAAGCGCTGTGAATAAGGGTCAATTAGAGCT 960
Qy 108 ----- 108
Db 961 ACCCTTCTTCTTCTATGTGTGAGTGTCTCAGCTTCCAGGTGACAGAGCCATCCCTGGG 1020
Qy 108 ----- 108
Db 1021 CGTCCAGAGAGACCTTGCTCCCTTTTGTGAGGGGCGCGCCCTCATTTGACTGTTCC 1080
Qy 108 ----- 108
Db 1081 ACCTTTGTGCTGGCTGTGGAGATTTTCTCTGAGATCTGCTTTAGAGCATGCT 1140
Qy 108 ----- 108
Db 1141 CAGGCCATTTCAAGTGAAGAGATGGGCTTCTGTCCGAGAGGGGTGTGTCAGCCTT 1200
Qy 108 ----- 108
Db 1201 TTCTGCTCTTACAGACTCAGGCTCCATTTGCCAGTATTTCTACAGCATTTACAGCAGG 1260
Qy 108 ----- 108
Db 1261 GGGAGAGGTGTCACTGTACCAAGTCTCTGGGGCTGTGGGGTTTGAAGAGAGCTCCC 1320
Qy 108 ----- 108
Db 1321 AGAAAGAGTTGTAGTATCATACAGAGCTGTCAAGGCTTGGCAGTTTCTGAGCAGTT 1380
Qy 108 ----- 108
Db 1381 GCCTGTGAGTGTCCAGGTTGTGACTGTCTTACTAACCCCTGGAACAACATTAACTCC 1440
Qy 108 ----- 108
Db 1441 GTTGTGAGACTGAGGGGGCAGAGGTCAACACAGCTGGGCTTACGCCAGCTTTGCT 1500
Qy 108 ----- 108
Db 1501 CTTAGCTTCGCACTGTCTCTGCTCCAGCCGTCTGACAGTGTGCTCTGAGGAGTGC 1560

Qy 108 ----- 108
Db 1561 ACACCCCTTCGAGGCTGAAGGCTTTGTCAAAGACATTGATCTCGAGGAAGCAGCCGGTG 1620
Qy 108 ----- 108
Db 1621 CCTTGCAGTGCTGGGCCAGAGCAGGACTGACATGTGGAAGCTCCATCACAGAACC 1680
Qy 108 ----- 108
Db 1681 TGCAGTGAAGACAGCAAGCTGCTGCTGTGATTAAATAGAGGACATTTGGCACCACTGAC 1740
Qy 109 ----- AlaPheAspValVal 113
Db 1741 AGTGGTGTTTGAACACAGCCTTTGGCCCTGCTGTGTGCTCCCTTAGGCTTCGATGTGGTG 1800
Qy 114 GluArgSerPheLeuGluSerIleAspAlaLeuAlaGluLysAlaSerLeuGlnSer 133
Db 1801 GAGAGGAGCTTCCTGGAGTCCATTACGACGCCTTGGCTGAGAAAGCAAGCCTCCAGTCG 1860
Qy 134 GlnLeuProGlu----- 137
Db 1861 CAATTGCCAGA-GGTAAATTTCCCGAGCCGACACCCAGGGAGTCAAGTCCAGGCCAGCT 1919
Qy 137 ----- 137
Db 1920 TTGCAAGGAGCATGGACTCATCTACTTTCTTGACATTACTGGGCCAGAGCAACAGGCGTT 1979
Qy 137 ----- 137
Db 1980 AGGAGCAGTTCTGTATGGTGACACTGGTGTGGGCCACAGGTGAGGGACTTCGCTGTCT 2039
Qy 137 ----- 137
Db 2040 CTGCTGTGGCAGCTGCATGCTAAAGCTCAGCAGTGTCTGGGCAGTGGTGACATGAGGA 2099
Qy 137 ----- 137
Db 2100 AGGGCTTTTCCAGTCTTCGTCCTGAATGGTCCCGCCCTTGTCAACCACCTGCTCAC 2159
Qy 137 ----- 137
Db 2160 ATCGCTACTGCCACCTGCAGCGCTTACCTCTGTGCTGTGAGTTGCGTATCTCCAGCTT 2219
Qy 137 ----- 137
Db 2220 TCTGGAAGAACCTTAGCCTGGAGCTGTAGCGAAGCTTCATCTTGGCTGTGCCATGGACC 2279
Qy 137 ----- 137
Db 2280 ATCTGTCTGTCTCCCCGGGCCCTCGGATCTTTTGCTCTTAAGAGCAAGGAGGCGCG 2339
Qy 137 ----- 137
Db 2340 GCATGTGGCTCAGGCTGTAAATCCAGCACCTTGGGAGGCCGAGGGCGGCGGATCACAA 2399
Qy 137 ----- 137
Db 2400 GGTCAAGAGATCAAGACCATCTGGCTAACATGGTGAAACCCGCTCTCTACTAAAAATAC 2459
Qy 137 ----- 137
Db 2460 AAAAAATTAGCTGGGGTGGTGGGGCGCCTTGTAGTCCAGCTACTCTCGAGGCTGAA 2519
Qy 137 ----- 137
Db 2520 GCAGGAGATGGTGTGAACCTGGGAGGGAGCTTGTAGTGAGCCGAGATCATGCCACTG 2579
Qy 137 ----- 137
Db 2580 CACTGTAGCCTGGGGGACACAGCGAGCTCCATCTCAAAAAAAAAAAAAAGAGTAAGGG 2639
Qy 137 ----- 137

Db 2640 AGTGGAGAGGGCTAGAAAGATGGTCCAGAGTTCTGTCTGCTGCCCTGACCCCTCTGTGATGG 2699
Qy 138 ----- GlyValProGlnHisGlnLeuProGlnIleValLeuGluArgLeu 155
Db 2700 TTGTAGGGAGTCCCTCAGCACCAGCTGCTCTCTCAGTATCAGAAGATCCTTTGAGAGACTC 2759
Qy 156 LysThrLeuGluArgGluIleSerGlyGlyAlaMetAlaValValAlaValLeuLeuAsn 175
Db 2760 AAGACGTTAGAGGGGAAATTTCCGGAGGGCCATGCCCGTTGTGGCGGCTCTTCTCAAC 2819
Qy 176 AsnLysLeuTyxValAlaAsnVal----- 183
Db 2820 AACAACTCTACGTCCGCAATGTCCGTGAGCCCTCTCTGCCAGGGCAGGGAGACTG 2879
Qy 183 ----- 183
Db 2880 GGGAGAGGTCAGCCACAGGGGTCCGTGCAITTAATTTGACAAATCTCTTTCCAGACACTTCA 2939
Qy 183 ----- 183
Db 2940 CGCACTTTAAACCCAGGGTCTCTGAGACCTTGGGTATGTCCCTCTCCACAGTGAAGCC 2999
Qy 183 ----- 183
Db 3000 TCAGTCCCAACTGGAAGGAGAAAGACGGATGGGAGACAGAGTGTCTGGCCCTTTAGTC 3059
Qy 183 ----- 183
Db 3060 CCTATTCTCTTCTTAACCTCATTTCTGGACGAGTCGTATCCCATTTCTGGGTGGCTTTGGGG 3119
Qy 183 ----- 183
Db 3120 GCGCGTCTTCTGGAAGAGTTTACCTGGCCATGAATAACCAACAGTATCCCATGAACA 3179
Qy 183 ----- 183
Db 3180 GCTGCAGATGGGCTAGGTGACAGGGACATTTGGGTTTGTGAGAACACAGACAGAGTGC 3239
Qy 183 ----- 183
Db 3240 AGSTTTTCAGTAGAAAGGACTCTGTAGAGACCTTCTGTATGATGTCGCTTTTAAATAC 3299
Qy 183 ----- 183
Db 3300 TCCAGTGGAAAAATTTCCACAATATAGACAATAGAGTGAATTAATGAACATTC 3359
Qy 183 ----- 183
Db 3360 ATGGCCCATCCCAATCCAGAGTTTATCAACTGTGGCCAGCCTCTTTGTCCCCACGT 3419
Qy 183 ----- 183
Db 3420 TCTCTATTTTGAAGCAAAATGCCAGACACTGTATCATGTTATCTGTAAATATTTCACTCTG 3479
Qy 183 ----- 183
Db 3480 TATCTCTAGAGAAGAGTCTGTTTATTAAAGATCATTAATCCTATGATTACACTGAAAAAG 3539
Qy 183 ----- 183
Db 3540 TTAACCCCAATTTTTCGTATCACAGCTGTCCCTAACTAGTGACTCACCCCGTTGGTCT 3599
Qy 183 ----- 183
Db 3600 GAGCCTGTTTGGCCCATTTTCAGGATTTCCATGTGTGAAATGCGCTTTTCCCTCTCT 3659
Qy 183 ----- 183
Db 3660 GCCTTCCCGGTATGCCCTATTCTCTCTGTGTGTAGTCTTTGTGCTAGCTGTTCACATTCT 3719
Qy 184 ----- GlyThrAsnArgAlaLeuLeuCysLysSerThrValAspGlyLeuGlnValThr 201
|||||


```

Db 3720 GCCACAGGTACAAACCGTGCATTTTATGCAATGCACAGTGGATGGTGACAGGTGACA 3779
Qy 202 G|H|L|E|U|S|N|V|A|L|A|P|H|S|T|H|T|R|G|I|U|A|N|G|I|U|A|S|P|I|U|L|E|U|P|H|E|A|I|S|E|R|G|I|U|A 221
    |||||
Db 3780 CAGCTGAACGTGACACACACACAGAAAGAGATGAGTCTCTCCGCTTTCGACAGCTG 3839
Qy 221 ----- 221
Db 3840 GGTGAGTGGGGAGAGTGGAGCGGAGAGCTGATCCCATGGGCTCACTTGGCCCTGCTT 3899
Qy 221 ----- 221
Db 3900 TGTGTGTGGGTAGAGAGCGGTGTGTAGAGGGGCTGTGATCTTGGCTCCCAAGCCAGC 3959
Qy 221 ----- 221
Db 3960 CTGCTGGGGTTCATTCGCCAGACCTGCCGCTTACTGTGTGTGATTTGAACCTTAAGTTT 4019
Qy 221 ----- 221
Db 4020 CTGTGCTCAGATCCCGCGCTGTGTATAGAGATTAATAGTAACCTAAGCTCGGGATGTAG 4079
Qy 221 ----- 221
Db 4080 GAGGATTCAGTTAGTGCATGTGAATGCTTCTGTGAGTGCCTGGCAACAGAAAGACCTT 4139
Qy 221 ----- 221
Db 4140 ACTGTTATGTGTGAGACTGAAAGAGGCCAAAGAGTCCAGAGAGCCAGCTGCTGTA 4199
Qy 221 ----- 221
Db 4200 GCTGCTCCTCCCTCCCAAGTGAAGCTCAAGGAGATGTGTGGGACACGSGATGGGACAGG 4259
Qy 221 ----- 221
Db 4260 AGACGGCAGCAGACGCTGCTCGTGCAGGTGTGCTGTGAGAGACGGGCTGTGAGTGTAG 4319
Qy 221 ----- 221
Db 4320 GTGTGTATAGGCGGTGGGACTGAGACACAGGAGCTTTGGGTAGCTGCTCCAGC 4379
Qy 221 ----- 221
Db 4380 GTGCCCTGTGTGTCTTCAATTCATTCAGTGGGTCTTATGCTTCTTCCCATGAC 4439
Qy 221 ----- 221
Db 4440 TGTGTCTGTCCCTCTTTTGTCTTGTGTGAACAAGAGAGATTTGCACT 4499
Qy 222 ----- 222
Db 4500 GTTTCCTCCCTAGGCTTGATGTGTGAAGATCAAGAGTGGGATCATCTGTGGGA 4559
Qy 223 n|g|l|u|s|e|r|h|a|p|h|a|l|e|g|y|a|s|p|i|u|l|e|u|p|h|e|a|i|s|e|r|g|i|u|a 257
    |||||
Db 4560 GGAAGACACCGGCGGATTCAGAGTTAAATATGCTTAACGAGATTTGACCT 4619
Qy 257 u|l|e|u|s|e|r|-----|A|a|l|a|y|s|e|r|y|s|p|r|o|i|e|i|l|e|a|g|l|u|p|r|o|i|l|h|i|s|e|l|y|a|l 275
    |||||
Db 4620 TCTCAGTAGTGCAGCCGAGCTGCTGCTGCTTGAAGAACAGAA----- 4668
Qy 275 a|g|i|n|p|r|o|l|e|u|a|s|p|i|u|l|e|u|p|h|e|a|i|s|e|r|g|i|u|a|l|e|u|p|h|e|a|i|s|e|r|g|i|u|a 295
    |||||
Db 4669 -CGTCTTGGGGGCGCAAGATGGAGATTTATGTGCTG-----GGAGCTCGAGGCTG 4721
Qy 295 l|a|l|e|u|g|l|u|a|l|h|i|s|e|l|y|-----|P|r|o|i|g|i|n|a|a|e|n|g|i|n|g|i|l|e|a|l| 310
    |||||
Db 4722 CAGTAGCCGTGATCATCTCACTGCATCCAGCTCGGGCAACAGAGACCAAGACCTGTCTC 4781
Qy 311 -----|A|l|e|t|i|l|e|a|s|p|h|r|g|i|u|p|h|e|a|i|l|y|s|i|n|t|h|s|e|r|l|e|u|a|p|h|a|l|a|l|a|g 328
    |||||
Db 4782 AAAAAAAAAAAAAAGGTCCAGAAAGCTTGGGGCAGAGCC---AGACTGTAGGCCCCAGTGA 4838

```

```

Qy 328 l|n|a|l|e|a|l|a|s|p|a|r|v|a|i|l|e|a|r|g|i|l|h|i|s|e|r|a|p|h|r|p|h|e|a|i|l|e|a|r|g|i|l|u|a 348
    |||||
Db 4839 CACATGTGCAGTCAAAATTAATTACAGAGCCAGGCAAACTTTCTCTC---CCAGAGA 4895
Qy 348 r|g|-----|A|a|r|g|h|e|c|y|s|p|r|o|a|r|g|h|i|s|g|i|a|s|p|h|e|r|l|e|u|e|u|a|l|a|r|a|s|n|p 365
    |||||
Db 4896 GATTGTTTGGGCGAGGTGAGCCAGAGCCAGGAGCTTGGGTACTGTAGAAACAAAGG 4955
Qy 365 n|e|g|i|l|y|r|p|r|o|l|e|u|g|i|l|u|e|t|-----|S|e|r|g|i|n|p|r|o|h|r|p|r|o|-----|S 378
    |||||
Db 4936 GTGAGTTGAGCTTGCCCAAGCTGAGGGCAGCACTAAGTCCCAAGAGAGGAGGCGGA 5015
Qy 378 e|r|p|r|o|a|l|p|r|o|a|l|a|l|g|i|l|y|a|r|g|v|a|i|l|y|r|p|r|o|v|a|i|s|e|r|-----|V|a|l|p|r|o|t|r|s|e|r|s 396
    |||||
Db 5016 AGCCAGATTGGCATGGATGGACAGACAGGCTTGTGTGCTGCTGCCATGGCTGC 5075
Qy 396 e|r|a|l|a|n|s|e|r|t|h|s|e|r|v|a|i|l|h|i|s|e|r|l|e|u|e|r|v|a|i|e|t|p|r|o|s|e|r|g|i|n|----- 414
    |||||
Db 5076 AGCAGAG-----TGTGAGATGAATGTGTATTTCCCTTGGAGCA 5117
Qy 415 -----|G|y|i|l|e|u|e|r|v|a|n|g|i|l|a|h|i|s|e|r|a 424
    |||||
Db 5118 CCTCTGTGACATCACCCAGCGCTCATGAGCAGAAATGGGCGATCATCATCACAGG 5177
Qy 424 l|a|s|e|r|t|h|l|e|u|a|s|p|i|u|l|a|t|h|r|p|r|o|h|i|l|e|u|r|h|a|s|n|g|i|n|s|e|r|p|r|o|h|i|l|e|u|r|h|e|u|g 444
    |||||
Db 5178 CCAGCCCGTGCAGAGCTGTGATGACCTTTTATGATTCATTCCTACTCAGGCCCCCT 5237
Qy 444 l|n|s|e|r|t|h|a|s|n|t|h|i|s|-----|T|h|i|n|g|i|n|s|e|r|s|e|r|s|e|r|s|e|r|a|s|p|i|o|l|y|i|l|e|u|p|h|e|r|g 463
    |||||
Db 5238 GAGGAGCTACTGCCCATCACCCAGAGGCATTTTCATTAAGAGAAACAGAGCATGAGAGA 5297
Qy 464 s|e|r|a|r|g|p|r|o|a|l|h|i|s|e|r|l|e|u|p|r|o|p|r|o|g|i|y 473
    |||||
Db 5298 CATCAGCAGCCAC-----CCTGGT 5318
    |||||

RESULT 7
US-09-925-300-330
; Sequence 330, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 330
; LENGTH: 696
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (643)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (657)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (685)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-300-330

Alignment Scores:
Pred. No.: 2,31e-29 Length: 696
Score: 409.00 Matches: 78

```

Percent Similarity:	98.75%	Conservative:	1
Best Local Similarity:	97.50%	Mismatches:	1
Query Match:	15.85%	Indels:	0
DB:	10	Gaps:	0

US-09-830-144-4 (1-504) x US-09-925-300-330 (1-696)

Qy	425	SerThrLeuAspGluAlaThrProThrLeuThrAsnGlnSerProThrLeuThrLeuGln	444
	:	:	:
Db	13	GCACCTCTGGACGAAGACACCCACCCCTCACCAACCAAGCCGACCTTAACCTGCAG	72
Qy	445	SerThrAsnThrHisThrGlnSerSerSerSerSerAspGlyGlyLeuPheArgSer	464
	:	:	:
Db	73	TCCACCAACACGACACGACGAGGAGCTCCAGCTCTRACGGAGGCGCTCTCCGCTCC	132
Qy	465	ArgProAlaHisSerLeuProProGlyGluAspGlyArgValGluProTyrValAspPhe	484
	:	:	:
Db	133	CGGCCCGCCCACTCGCTCCCGCTGGCGAGGACGGTCGTGTGAGCCCTATGTGGACTTT	192
Qy	485	AlaGluPheTyrArgLeuTrpSerValAspHisGlyGluGlnSerValThrAlaPro	504
	:	:	:
Db	193	GCTGAGTTTTACCGCTCTGGAGCGTGGACCAATGGCGAGGAGCGTGGTACACGACCG	252

RESULT 8

```

US-09-864-761-17553
; Sequence 17553, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeonics-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29

```

```

? NUMBER OF SEQ ID NOS: 49117
?
? SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
? SEQ ID NO 17553
? LENGTH: 211
? TYPE: DNA
? ORGANISM: Homo sapiens
?
? FEATURE:
?
? OTHER INFORMATION: MAP TO Z83845.14
? OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
?
? OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.77
?
? OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1
?
? OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.8
?
? OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3
?
? OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.73
?
? OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
?
? OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.3
?
? OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.78
?
? OTHER INFORMATION: EXPRESSED IN HBu100, SIGNAL = 1.4
?
? OTHER INFORMATION: NT HIT: g15174702, EVALUATE 1.00e-115
?
? OTHER INFORMATION: SWISSPROT HIT: Q15750, EVALUATE 5.00e-36
?
? OTHER INFORMATION: EST_HUMAN HIT: BE898567.1, EVALUATE 1.00e-115
?
? US-09-864-761-17553

```

Alignment Scores:

Alignment Scores:					
Pred. No.:	5.14e-26	Length:	211		
Score:	365.00	Matches:	70		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	14.15%	Indels:	0		
DB:	10	Gaps:	0		
US-09-830-144-4 (1-504) x US-09-864-761-17553 (1-211)					
Qy	312	MetIleAspThrGluPheAlaLysGlnThrSerLeuAspAlaValAlaGlnAlaVal	331		
Db	1	ATGATTGACACTGAGTTGCCAGACGACCTCCCTGGACGCAGTGGCCACAGCCGCTGGTG	60		
Qy	332	AspArgValLysArgIleHisSerAspThrPheAlaSerGlyGlyGluValGAlaAarPhe	351		
Db	61	GACC GGTTGAAGCGCATCCACAGCAGACACTTCGCCAGTGGTGGGGAGCGTGCCAGGTTTC	120		
Qy	352	CysProArgHisGluAspMetThrLeuLeuValIArgAsnPheGlyTyrProLeuGlyGlu	371		
Db	121	TGCCCCCGCAGCAGGACATGACCTTGCTAGTGAGGAACCTTTGGCTACCCCTGGGGCAA	180		
Qy	372	MetSerGlnPnroThrProSerProAlaPro	381		
Db	181	ATGAGCCAGCCACACCGAGCCAGCCCA	210		

RESULT 9

```

US-09-864-761-769
; Sequence 769, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Rank, Sharron G.
; APPLICANT: Penn, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27

```

```

; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 769
; LENGTH: 455
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO 283845.14
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.77
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.73
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.78
; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 1.4
; US-09-864-761-769
Alignment Scores:
Pred. No.: 1,01e-21 length: 455
Score: 323.50 Matches: 73
Percent Similarity: 79.00% Conservative: 6
Best Local Similarity: 73.00% Mismatches: 12
Query Match: 12.54% Indels: 9
DB: 10 Gaps: 3
US-09-830-144-4 (1-504) x US-09-864-761-769 (1-455)
Cy 274 G|y|a|a|g|n|p|o|l|e|u|a|s|p|g|y|a|t|h|t|g|l|y|p|h|e|u|v|a|l|e|u|m|e|t|e|r|g|l|u|g|l|y|e|u|t|r 293
Db 168 G|G|C|T|G|T|G|A|C|C|A|C|T|G|A|T|-----CTC|C|T|G|A|T|T|T|A|G|C|T|C|C|A|G|A|T|T|A|G|G|C 215
Cy 294 Lys|Ala|Leu|Gln|Ala|Ala|His|Gly|-----Pro|G|L|G|L|Ala|-----Aen|G|L|G|L|u 308
Db 216 C|A|T|G|C|C|C|C|C|C|C|G|C|C|G|T|C|C|T|T|A|C|A|G|G|T|C|T|C|C|T|A|C|C|C|C|C|C|C|C|G|G|A|G 275
Cy 309 I|l|e|a|l|a|Met|l|e|a|s|p|T|h|t|G|L|P|h|e|a|l|y|s|G|L|n|h|t|e|r|e|u|s|p|a|l|a|v|a|l|a|G|L|n 328
Db 276 A|T|T|G|C|T|G|G|A|G|A|T|T|G|A|C|A|G|A|T|T|T|G|C|C|A|G|A|C|C|T|C|C|G|A|C|G|A|G|G|G|C|C|A|G 335
Cy 329 A|A|a|V|a|l|a|s|P|a|T|G|V|a|l|y|e|a|G|l|e|h|s|e|s|e|s|p|h|t|h|e|a|l|a|s|e|r|g|l|y|G|L|u|A|T|G 348
Db 336 G|C|C|G|C|G|T|G|G|A|C|C|G|G|T|A|A|C|G|C|A|T|C|C|A|G|C|A|C|C|T|T|G|C|G|C|A|T|G|T|G|G|A|G|C|G|T 395
Cy 349 A|A|a|T|p|h|e|C|y|s|P|r|o|A|T|G|H|s|G|L|u|s|M|e|t|h|e|r|e|u|v|a|l|a|T|g|a|n|p|h|e|G|l|Y|T|r|P|ro 368
```

```

Db 396 G|C|C|A|G|T|T|G|C|C|C|C|G|C|A|C|A|G|A|C|A|T|G|A|C|C|T|G|T|A|G|A|A|C|T|T|G|G|T|A|C|C|G 455
RESULT 10
US-09-864-761-770
; Sequence 770, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmtca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 770
; LENGTH: 467
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO 283845.14
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 10
; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 5.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.7
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 8.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 8.9
; US-09-864-761-770
```

Alignment Scores:
Pred. No.: 6,64e-17 Length: 467
Score: 271.50 Matches: 72
Percent Similarity: 56.74% Conservative: 8
Best Local Similarity: 51.06% Mismatches: 23
Query Match: 10.52% Indels: 38
DB: 10 Gaps: 5

US-09-830-144-4 (1-504) x US-09-864-761-770 (1-467)

QY 1 MetAlaGlnArgArgSerLeuGlnSerGluGlnProSer-----16
Db 56 GTGGCCCGTGAGAGGTGGCTCTGCTGTCTTGCACAGGCGCTGCTGTGATGGGGTAG 115
QY 17 -----TTPThrAspAspLeuProLeuGlnSerGlyValGlySerAlaSer 33
Db 116 CGTGACATGGGAGAGGTATCCCAATGTCTAT-----ACCCAGAT 160
QY 34 AsnArgSerTyrSerAlaAspGly-----LysGlyThrGluSer-----46
Db 161 GAAATGATGGCTAAAGCAGGGGGACCCAGAGGGCCCTGAAAGCTGCAGCTGTGCTGCT 220
QY 47 -----HisProGluAspSerTrpLeuLys 55
Db 221 TTAGTCTCCCCAATCTTTCCCTTCTCCCTCCACTCCGTGAGACCCCTGCTCTCAG 280
QY 56 -----Phe-ArgSerGluAsnAsnCysPheLeuTyrGlyValPh 68
Db 281 CTTCCCTCTGCCCTCTCCCTCTTCCAGAGTGAGAACAACTGCTTCTGTATGGGGTCTT 340
QY 68 eAnGlyTyrAspGlyAsnArgValThrAsnPheValAlaGlnArgLeuSerAlaGluLe 88
Db 341 CAACGGCTATGATGGCAACCCAGTGACCACTTCGTGGCCAGCGGTGTCCCGAGCT 400
QY 88 uLeuLeuGlyGlnLeuAsnAlaGluHisAlaGluAlaAspValArgValLeuLeuGln 108
Db 401 CTTGCTGGGCCAGCTGAATGCCAGACCGCCGAGCGGATGTGGGGCTGTGCTGTGCA 460
QY 108 n 108
Db 461 G 461

RESULT 11
US-09-864-761-17554
; Sequence 17554, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Penn, Sharron G.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Acomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 17554
; LENGTH: 194
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO Z83845.14
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 10
; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 5.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.7
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 8.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.4
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 8.9
; OTHER INFORMATION: NT HIT: q11418147, EVALUE 3.00e-82
; OTHER INFORMATION: SWISSPROT HIT: Q15750, EVALUE 8.00e-24
; OTHER INFORMATION: EST_HUMAN HIT: AL118967.1, EVALUE 5.00e-82
US-09-864-761-17554

Alignment Scores:
Pred. No.: 1.84e-16 Length: 194
Score: 261.00 Matches: 51
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.12% Indels: 0
DB: 10 Gaps: 0

US-09-830-144-4 (1-504) x US-09-864-761-17554 (1-194)

QY 58 SerGluAsnAsnCysPheLeuTyrGlyValPheAsnGlyTyrAspGlyAsnArgValThr 77
Db 2 AGTGAGACAACTGCTTCTGTATGGGGTCTTCAACGGCTATGATGGCAACCGAGTGACC 61
QY 78 AsnPheValAlaGlnArgLeuSerAlaGluLeuLeuGlyGlnLeuAsnAlaGluHis 97
Db 62 AACTTCGTGGCCAGCGCTCTCCGACAGCTCTGCTGGCCAGCTGAATGCCGAGCAC 121
QY 98 AlaGluAlaAspValArgValLeuGln 108
Db 122 GCCGAGGCCGATGTCGGCGTGTCTGCTGCAG 154

RESULT 12
US-09-938-842A-2158
; Sequence 2158, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong


```

Db      582 TTCTAGATCACAAACAGATCTTGAAGCTGAGAAAGAAAGATA-----TT 629
Qy      223 UASPAIAGLYLVYIILEYSGINVAIGLYILEYSGIN---GlusterThargar 242
Db      630 GAAAGCTGGGGCTTATATACATGACGGGCGAGTCAATGAAGCTTAATATATACGAGC 669
Qy      242 gIleGIAspTYrLYsValIySTyrgLYrThrasPIleAspLeuSerAlaIaly 262
Db      690 TATCGGATCATGGAATTCAGACAG-----AATAAGTTTGGCATCTGA 724
Qy      262 sSerLYsProIIleIAlAGIuProGIuIleHISGLYAlAGInProleuAspGIYAlTh 282
Db      735 AAGACAAATAGTACCGCTAGTCCAGATGTATACCTGTGAAGCTGTGTGAT--GA 791
Qy      282 rGIYpHeLeuValIleuMeSerGIuGIyLeuTYrLYsAlaAGIuAlaIaHISGLYPr 302
Db      792 TGATTTCTGTTCTTGGCTGCGATGGAATTTGGGATTCATG----- 834
Qy      302 oGIYGIAlAsnGIuIleAlAlaMeTlleAspThrGIuPHeAlaLYSGInThrSe 322
Db      835 -----ACAAGCCAACACTCGTTGATTTTCATACATGAACATTGAATTCAGAGACCA 887
Qy      322 rLeuAspAlaValAlaGIAlaValaValaAspArGValysArGIleHISserAspThrPh 342
Db      888 ACTCGCGTGTATATGAAAAGTTCTCGATGATGTCTGCGCTCCAAACT----- 939
Qy      342 eAlAsErGIYGLYluArGIAlaArpHeCySProArGISeGIuAspMeTThrLeuIeUa 362
Db      940 -----TCAGGGGTGGAAGGCTGT-----GATAACATGACCATGATATT 977
Qy      362 lArGAspHeGIYTYrProleuGIYluMeSerGIuProThrProserProAlaProAl 382
Db      978 GCTTCGATTCAAGAACCTCTCTCCATCAGAGACCAACCAAGACAGACCGAGGC 1037
Qy      382 aAlAGLYLYArGValTYrProValSer 391
Db      1038 AGAAGAAACCAAGATGAGCCGAGCTCA 1065

```

RESULT 15

US-09-770-445-6
 Sequence 6, Application US/09770445
 Patent No. US20020023281A1

GENERAL INFORMATION:
 APPLICANT: Gorlach, Jörn

APPLICANT: An, Yong-Qiang

APPLICANT: Hamilton, Carol M.

APPLICANT: Price, Jennifer L.

APPLICANT: Raines, Tracy M.

APPLICANT: Yu, Yang

APPLICANT: Rameaka, Joshua G.

APPLICANT: Page, Amy

APPLICANT: Matthew, Abraham V.

APPLICANT: Ledford, Brooke L.

APPLICANT: Woessner, Jeffrey P.

APPLICANT: Haas, William David

APPLICANT: Garcia, Carlos A.

APPLICANT: Krickler, Maja

APPLICANT: Slader, Ted

APPLICANT: Davis, Keith R.

APPLICANT: Allen, Keith

APPLICANT: Hoffman, Neil

APPLICANT: Hubban, Patrick

TITLE OF INVENTION: Expressed Sequences of Arabidopsis

TITLE OF INVENTION: thaliana

FILE REFERENCE: 2023US (PARA-012PRV)

CURRENT APPLICATION NUMBER: US/09/770,445

CURRENT FILING DATE: 2001-01-26

PRIOR APPLICATION NUMBER: US 60/178,472

PRIOR FILING DATE: 2000-01-27

NUMBER OF SEQ ID NOS: 999

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 6

LENGTH: 1496

```

; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1496)
; OTHER INFORMATION: n = A,T,C or G
US-09-770-445-6

Alignment Scores:
Pred. No.: 2,61e-08 Length: 1496
Score: 166.00 Matches: 97
Percent Similarity: 42.39% Conservative: 70
Best Local Similarity: 24.62% Mismatches: 161
Query Match: 7.21% Indels: 67
DB: Gaps: 17

US-09-830-144-4 (1-504) x US-09-770-445-6 (1-1496)
Qy      21 leuProleuCyS-HISleuSerGIY-----ValGIYserAlase 33
Db      131 CTTCCTTTTGTTCATTTCTTACCCCATATATCATGGGTACATACCTTAAGTTCTCCGAA 190
Qy      33 rAsnArGserTYrSeAlaAspGIYsGLYrThrGIuSerHISProGIuAspSerTr 53
Db      191 AACTGAAAGTTATCGAAGATGTGAGAAATGATTAAGCTCAGATTGGTTATGCTAT 250
Qy      53 pLeuLYsPheArGser-----GUAAs 61
Db      251 GCAGAGTTGGGGCTTACCATGGAAGACCGCATGCTGCATTTCTTGATTTGATGATAA 310
Qy      61 nCySPheLeuTYrGIYAlPheAsnGIYTYrAspGIYsArGValThraspHeValAl 81
Db      311 GACATCGTCTTGTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 370
Qy      81 aGIArGleuSerAlaGIuLeuLeuLeuGIYIn---LeuAsnAlaGIuHISAlaGIuAl 100
Db      371 CAAGATCTACACCGAGGGTATTCAGTATGAGCGTATTAATTCGAGACGTCGAAAC 430
Qy      100 aAspValArGArGValLeuLeuGIuAlaPheAspValAl-----GIuArGserPheLe 118
Db      431 ATCTCTTGAAGAGCATTTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 490
Qy      118 uGIuSerIleAspAspAlaGIuLeuAlaGIuLYsAlaSerIleuGIuSerGIuLeuProGIuGI 138
Db      491 AGAGTTAGCT-----GTACTTGGCGACAGATGAACAATTTACGGCGATGATTTGAAGG 544
Qy      138 Y-ValProGIuHISGIIn-----LeuProGIuInTYrGIuLYsIleuGIuArGleu 156
Db      545 ATTTATATGCTCACCAAGAGCGGTGACACCAATTAACCAACCGATAGTGGCTCTTGA 604
Qy      156 YS-----ThreGIuArGIuIleSerGIYAlaMeTlleAlaValAlaVal 172
Db      605 AGATGCTTCATTTCTGATTTCAAGGACCTACCTCGGGGTGCAACA-CCGTGTAGCTC 663
Qy      172 aLeuLeuAsnAsnLYsLeuTYrValAlaAsnValGIYThrasnArGAlaLeuLeuCYSL 192
Db      664 TTATTAAAGATPAAGAACCTCTTTGTTCGAAATGCCGTACCTCAGCTGTGTGATATCAAA 723
Qy      192 ySerThrValAspGIYleuGIuAlThreGIuLeuAsnValAspHISThrThrGIuAsnG 212
Db      724 GAAAGAT-----CAGGCTTACATTTCTTAAATACCAAGCCTGATCTTG 771
Qy      212 lUAspGIuLeuPheArGleuSerGIuLeuGIYleuAspAlaGIYsIleYsGIuValG 232
Db      772 AAGTTGAAGAAAGAAAGATA-----TTGAACTGGTGGCTTTATTCACGCTG 819
Qy      232 lYIleIleCYsGLYIn---GlusterThrArGArGIleGIYsPyrTYsValIySTyrg 251
Db      820 GGAGAAATCAATGAAGCTTGAATCTGACAAAGACCTTGTGTATATGAGTTCAAGCAG- 878
Qy      251 lYTYrThrasPIleAspLeuSerAlaAlaLYsSerIySProIIleIleAlaGIuProG 271
Db      879 -----AATAAGTTTTCACATCTGAAAGCAATGTTACTGCTATCCAG 924

```

```
Qy 271 luileHisGlyAlaGlnProLeuAspGlyValThrGlyPheLeuValIleuMetSerGluG 291
Db 925 ATATAAACACTATTGACCTATGTGATGAT---GATGACTTTCTTGTGTTNNNNNNNTN 981
Qy 291 lyLeuTyrlYsAlaLeuGluAlaAlaHisGlyProGlyGlnAlaAsnGlnIleAlaA 311
Db 982 NNANATGGGATTGTATG-----TCAGGCCAGGAAGTAGTTG 1017
Qy 311 laMetIleAspThrGluPheAlaLysGlnThrSerLeuAspAlaValAlaGlnAlaValV 331
Db 1018 ATTTTATCCATGAACAGTTAAATCTGAAACAAACTTTCAACAGTATGTCAAAAGGTTG 1077
Qy 331 alaAspArgValLysArgIleHisSerAspThrPheAlaSerGlyGlyGluArgAlaArgP 351
Db 1078 TTGATAGATGT-----TTGGCTCCAGATACACGCGACTGGTGAAGGT----- 1118
Qy 351 heCysProArgHisGluAspMetThrLeuLeuValArgAsnPheGlyTyProLeuGlyG 371
Db 1119 --TGT-----GATAATATGACCATCATCTTGGTTTCAGTTCAAGAAGCCTAACCCAT 1167
Qy 371 luMetSerGlnProThrProSer---ProAlaProAla 382
Db 1168 CTGAGACTGAACCAAGAGATTCCAAACCAAGACCAAGC 1205
```

Search completed: December 10, 2002, 02:09:31
Job time : 98.5983 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 10, 2002, 05:57:28 ; Search time 85 Seconds
(without alignments)
106.600 Million cell updates/sec

Title: US-09-830-144-4_COPY_437_504
Perfect score: 359
Sequence: 1 OSPRLTQSTWTHQSSSSS.....AEFRYSVDHGEQSVTAP 68

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A.Geneseq.101002:*

- 1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	359	100.0	504	18	AAW26706
2	359	100.0	504	18	AAW26707
3	359	100.0	504	20	AAW09541
4	359	100.0	504	21	AAAY91001
5	359	100.0	504	21	AAAY94450
6	359	100.0	513	20	AAAY09550
7	359	100.0	517	20	AAAY09546
8	352	98.1	84	21	AAAB5692
9	77	21.4	16	20	AAAY09549
10	70	19.5	22	22	ABW70838

11	69.5	19.4	199	22	AAG78804	Human HT4SG64, ser1
12	68.5	19.1	113	22	AAU47075	Protonibacterium
13	67.5	18.8	739	22	ABG16477	Novel human diageno
14	67.5	18.8	739	22	AAAG6173	Atrophin-1, interec
15	67.5	18.8	852	20	AAAG30948	Human E3 ubiquitin
16	66	18.4	16	20	AAAY09548	Human TAB1 peptide
17	64.5	18.0	116	20	AAU32515	Novel human secret
18	64.5	18.0	244	21	AAAG6525	Human gene 72-enco
19	63.5	17.7	187	21	AAAG5655	Arabidopsis thalia
20	63.5	17.7	219	21	AAAG5654	Arabidopsis thalia
21	63.5	17.7	240	19	AAAG75243	Fragment of human
22	63.5	17.7	241	22	ABG03626	Novel human diageno
23	63.5	17.7	243	21	AAAG35653	Arabidopsis thalia
24	63.5	17.7	335	19	AAAG75208	Human secreted pro
25	63.5	17.7	391	22	AAAG36176	Human DIB2 protein
26	63.5	17.7	824	22	AAAG20251	Human ADAM6, Homo
27	63.5	17.7	824	23	AAU86156	Human PPO1666, poly
28	63	17.5	66	22	AAW86430	Human immune/haema
29	63	17.5	2112	22	ABG60403	Drosophila melanog
30	62.5	17.4	92	21	AAAG2825	Human ORFX ORF2589
31	62.5	17.4	675	22	AAU50204	Protonibacterium
32	62.5	17.4	1945	22	ABG64947	Drosophila melanog
33	62	17.3	527	19	AAW48419	Amino acid sequenc
34	62	17.3	734	22	AAAB82316	Human immunoglobul
35	62	17.3	3583	22	ABG64814	Drosophila melanog
36	61	17.0	60	22	AAU66449	Protonibacterium
37	61	17.0	322	22	ABG10298	Human cDNA SEQ ID
38	61	17.0	1511	21	AAAG28182	Yeast transporter
39	60.5	16.9	854	20	AAAG30949	Murine E3 ubiquiti
40	60	16.7	1037	20	AAW67643	A serine/threonine
41	59	16.4	173	23	ABG55244	Lactococcus lactis
42	59	16.4	1046	22	AAE11775	Human kinase (PKIN
43	59	16.4	1327	22	ABG58539	Drosophila melanog
44	58.5	16.3	90	22	AAU56766	Protonibacterium
45	58.5	16.3	1093	16	AAAG6460	At-17 protein, Ho

ALIGNMENTS

RESULT 1	AAW26706	standard; Protein: 504 AA.
ID	AAW26706;	
AC	AAW26706;	
XX		
XX	14-APR-1998	(first entry)
DT		
XX		
XX	Human TAB1 (TAK1 binding protein).	
DE		
XX	TAB1, TAK1 binding protein; transforming growth factor-beta;	
KW	signal transduction; human.	
XX		
XX	Homo sapiens.	
OS		
XX		
XX	Key	Location/Qualifiers
FT	Misc-difference 52	/note= "variant has Arg as residue 52"
FT		
XX		
XX		
PN	EP803571-AA.	
PD		
XX	29-OCT-1997.	
XX		
XX	24-APR-1997;	97EP-0302808.
PF		
XX		
XX	20-NOV-1996;	96US-0752891.
XX	24-APR-1996;	96JP-0136282.
PR	28-OCT-1996;	96JP-0300856.
XX		
XX	(UENO/) UENO N.	
PA		
XX	Human prostate can	
XX	Human TAB1 peptide	
XX	Matsumoto K, Nishida E;	
PI		
XX		

DR WPI; 1997-515318/48.
 DR N-PSDB; AAT91175.
 XX
 PT DNA encoding TAK1 binding protein TAB1 - member of transforming
 PT growth factor beta receptor signal production pathway, which
 PT activates TAK-1 kinase activity upon binding
 XX
 PS Example 5; Page 17-19; 30pp; English.
 XX
 CC This protein comprises human TAB1, a novel member of the
 CC transforming growth factor-beta receptor signal transduction
 CC pathway, which activates TAK-1 kinase activity upon binding. Its
 CC amino acid sequence was deduced from a cDNA clone (see AAT91175)
 CC obtained from a kidney library; a variant TAB1 (see AAW26707) has
 CC Arg rather than Ser at amino acid position 52. Also claimed are:
 CC isolated DNA encoding a protein modified by a substitution,
 CC deletion and/or addition of 1 or more amino acids of the 504-residue
 CC TAB1 sequence; (2) DNA which can hybridise with the 1560 bp TAB1
 CC nucleic acid sequence; (3) isolated DNA encoding a protein
 CC comprising amino acids 21-579 or 437-504 of the 504 TAB1 sequence;
 CC (4) DNA encoding a fusion protein comprising an above protein or
 CC polypeptide; (5) expression vector comprising an above DNA; and
 CC the expression vector. Cells expressing TAB1 and TAK1 can be used
 CC to screen for TGF-beta signalling pathway inhibitors by contacting
 CC the cells with a test compound, and measuring the TAK1 kinase
 CC activity.
 XX
 SQ Sequence 504 AA;
 Query Match 100.0%; Score 359; DB 18; Length 504;
 Best Local Similarity 100.0%; Pred. No. 3.1e-37;
 Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QSPFTLTQSTNTHQTQSSSSDGLFRSRPAHSLPPGEGRVEPYVDFAEFYRLWSVDHG 60
 DB 437 QSPFTLTQSTNTHQTQSSSSDGLFRSRPAHSLPPGEGRVEPYVDFAEFYRLWSVDHG 496
 QY 61 EQSVVTAP 68
 DB 497 EQSVVTAP 504
 RESULT 2
 AAW26707
 XX ID AAW26707 standard; Protein; 504 AA.
 XX AC AAW26707;
 XX DT 14-APR-1998 (first entry)
 XX DE Human TAB1 (TAK1 binding protein).
 XX KW TAB1; TAK1 binding protein; transforming growth factor-beta;
 XX signal transduction; human.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 FT Misc-difference 52
 FT /note= "variant has Ser as residue 52"
 XX EP803571-A2.
 XX 29-OCT-1997.
 XX 24-APR-1997; 97EP-0302808.
 XX 20-NOV-1996; 96US-0752891.
 XX 24-APR-1996; 96JP-0126282.
 XX 28-OCT-1996; 96JP-0300856.
 XX (UENO/) UENO N.
 PA

XX Matsumoto K, Nishida E;
 XX WPI; 1997-515318/48.
 DR N-PSDB; AAT91178.
 XX
 PT DNA encoding TAK1 binding protein TAB1 - member of transforming
 PT growth factor beta receptor signal production pathway, which
 PT activates TAK-1 kinase activity upon binding
 XX
 PS Example 5; Page 19-21; 30pp; English.
 XX
 CC This protein comprises human TAB1, a novel member of the
 CC transforming growth factor-beta receptor signal transduction
 CC pathway, which activates TAK-1 kinase activity upon binding. Its
 CC amino acid sequence was deduced from a cDNA clone (see AAT91176)
 CC obtained from a kidney library; a variant TAB1 (see AAW26706) has
 CC Ser rather than Arg at amino acid position 52. Also claimed are:
 CC isolated DNA encoding a protein modified by a substitution,
 CC deletion and/or addition of 1 or more amino acids of the 504-residue
 CC TAB1 sequence; (2) DNA which can hybridise with the 1560 bp TAB1
 CC nucleic acid sequence; (3) isolated DNA encoding a protein
 CC comprising amino acids 21-579 or 437-504 of the 504 TAB1 sequence;
 CC (4) DNA encoding a fusion protein comprising an above protein or
 CC polypeptide; (5) expression vector comprising an above DNA; and
 CC (6) host cell, preferably a mammalian or yeast cell, transformed by
 CC the expression vector. Cells expressing TAB1 and TAK1 can be used
 CC to screen for TGF-beta signalling pathway inhibitors by contacting
 CC the cells with a test compound, and measuring the TAK1 kinase
 CC activity.
 XX
 SQ Sequence 504 AA;
 Query Match 100.0%; Score 359; DB 18; Length 504;
 Best Local Similarity 100.0%; Pred. No. 3.1e-37;
 Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QSPFTLTQSTNTHQTQSSSSDGLFRSRPAHSLPPGEGRVEPYVDFAEFYRLWSVDHG 60
 DB 437 QSPFTLTQSTNTHQTQSSSSDGLFRSRPAHSLPPGEGRVEPYVDFAEFYRLWSVDHG 496
 QY 61 EQSVVTAP 68
 DB 497 EQSVVTAP 504
 RESULT 3
 AAY09541
 XX ID AAY09541 standard; Protein; 504 AA.
 XX AC AAY09541;
 XX DT 21-JUL-1999 (first entry)
 XX DE Human TAB1 protein.
 XX KW Human; TAB1; TAK1; screening; inhibition; TGF-beta;
 XX transforming growth factor beta.
 XX OS Homo sapiens.
 XX WO9921010-A1.
 XX 29-APR-1999.
 XX 22-OCT-1998; 98WO-JP04796.
 XX 22-OCT-1997; 97JP-0290189.
 XX (CHUS) CHUGAI SEIYAKU KK.
 XX Ohtomo T, Ono K, Tsuchiya M;
 XX

DR WPI; 1999-312645/26.
 DR N-PSDB; AAX56278.
 PT Screening for TGF- beta inhibitory substances, which are useful as
 XX drugs for treatment of diseases relating to its disorder
 XX
 XX
 XX Claim 3; Page 147-149; 195pp; Japanese.
 PS
 CC A method has been developed for screening for substances which inhibit
 CC the binding of TAK1 polypeptide to TAB1 polypeptide. The method
 CC comprises: (a) contacting the polypeptide in the presence of a sample;
 CC and (b) detecting the amount of bound polypeptide, in which the sample
 CC can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming
 CC growth factor (TGF)-beta inhibitory substances can be used in drugs for
 CC indications e.g. as TGF-beta signal transduction inhibitors or
 CC activators, or extracellular matrix protein production enhancement
 CC inhibitors, or activators, or cell proliferation prevention inhibitors or
 CC activators, or monocyte migration inhibitors or activators, or
 CC physiological activity induction inhibitors or activators, or
 CC immunosuppression inhibitors or activators, or amyloid beta protein
 CC precipitation inhibitors or activators, and such substances can also be
 CC inhibitors of the TAK1 polypeptide function, particularly kinase
 CC activity. The present sequence represents human TAB1.
 CC
 XX
 SQ Sequence 504 AA;
 Query Match 100.0%; Score 359; DB 20; Length 504;
 Best Local Similarity 100.0%; Pred. No. 3.1e-37;
 Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QSPFLTIQSTNTHQTSSSSSDGGLFRSRPAHSILPFGEDGRVEPYVDFAEFYRLMSVDHG 60
 DB 437 QSPFLTIQSTNTHQTSSSSSDGGLFRSRPAHSILPFGEDGRVEPYVDFAEFYRLMSVDHG 496
 QY 61 EQSVVTAP 68
 DB 497 EQSVVTAP 504
 Db
 RESULT 4
 AAY91001
 ID AAY91001 standard; protein; 504 AA.
 XX
 AC AAY91001;
 XX
 DT 04-SEP-2000 (first entry)
 XX
 DE Human TAB-1 protein sequence SEQ ID NO:4.
 XX
 KW Human; TAK-1; TAB-1; mitogen activated protein kinase; MAPK;
 KW screening; signal transduction; inhibition; inflammatory cytokine;
 KW IL-1; interleukin 1; TNF; tumour necrosis factor; inflammation;
 KW antiinflammatory; suppression.
 XX
 OS Homo sapiens.
 OS
 PN WO200023610-A1.
 XX
 PD 27-APR-2000.
 XX
 PF 21-OCT-1999; 99WO-JP05817.
 XX
 PR 21-OCT-1998; 98JP-0299962.
 XX
 PA (CHUDS) CHUGAI SEIYAKU KK.
 XX
 PI Tsuchiya M, Ohtomo T, Sugamata Y, Matsumoto K;
 XX
 DR WPI; 2000-339707/29.
 DR
 DR N-PSDB; AAA39106.
 PT Method for screening inhibitors of TAK1 signal transduction for
 PT suppression of inflammatory cytokine production and use as

PT antiinflammatory agents -
 XX
 PS Disclosure; Page 90-94; 100pp; Japanese.
 XX
 CC The present invention describes a method for screening compounds for
 CC inhibition of inflammatory cytokine signal transduction by contacting
 CC the sample with TAK1 and its receptor TAB1 and selecting for inhibition
 CC of TAK1/TAB1 binding. Also described is a method for screening compounds
 CC for inhibition of inflammatory cytokine signal transduction in which the
 CC composition for the treatment of inflammatory disorders containing as
 CC active component an inflammatory cytokine signal transduction inhibitor.
 CC TAK1 is an essential component of the signalling process which results
 CC in release of inflammatory cytokines such as interleukin-1 (IL-1).
 CC IL-10, tumour necrosis factor (TNF) and IL-6. The methods can be used
 CC for the selection of effective antiinflammatory agents. The present
 CC sequence represents human TAB-1, which is used in the exemplification of
 CC the present invention.
 CC
 XX
 SQ Sequence 504 AA;
 Query Match 100.0%; Score 359; DB 21; Length 504;
 Best Local Similarity 100.0%; Pred. No. 3.1e-37;
 Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QSPFLTIQSTNTHQTSSSSSDGGLFRSRPAHSILPFGEDGRVEPYVDFAEFYRLMSVDHG 60
 DB 437 QSPFLTIQSTNTHQTSSSSSDGGLFRSRPAHSILPFGEDGRVEPYVDFAEFYRLMSVDHG 496
 QY 61 EQSVVTAP 68
 DB 497 EQSVVTAP 504
 Db
 RESULT 5
 AAY59450
 ID AAY59450 standard; protein; 504 AA.
 XX
 AC AAY59450;
 XX
 DT 24-MAR-2000 (first entry)
 XX
 DE Human TAB1 protein sequence.
 XX
 KW Human; TAB1; XIAP; X-linked inhibitor of apoptosis protein; TGF-beta;
 KW transforming growth factor-beta activated kinase 1; monocyte migration;
 KW TAK1 binding protein 1; extracellular matrix protein production;
 KW cell growth inhibitor; beta-amyloid protein deposition;
 KW immunosuppression; Transforming growth factor-beta.
 XX
 OS Homo sapiens.
 OS
 PN JP1326328-A.
 XX
 PD 26-NOV-1999.
 XX
 PF 13-MAY-1998; 98JP-0130378.
 XX
 PR 13-MAY-1998; 98JP-0130378.
 XX
 PA (MATS/) MATSUMOTO K.
 XX
 DR WPI; 2000-078337/07.
 DR
 DR N-PSDB; AAZ48861.
 PT Screening a substance which inhibits combination of the X-linked
 PT inhibitor of apoptosis protein -
 XX
 PS Claim 2; Page 25-26; 43pp; Japanese.
 XX
 CC This sequence represents the human TAB1 protein.
 CC The invention relates to a method for screening a substance inhibiting
 CC the formation of a complex between XIAP and TAB1, in which X-linked

CC inhibitor of apoptosis protein (XIAP), transforming growth factor-beta
 CC activated kinase 1(TAK1) binding protein 1(TAB1) and a substance to be
 CC tested are contacted with each other and then the presence or formation
 CC of a complex between XIAP and TAB1 is detected. The substance can be used
 CC as a drug for extracellular matrix protein production enhancement, cell
 CC growth inhibition, monocyte migration, physiologically active substance
 CC induction, immunosuppression, and beta-amyloid protein deposition. A
 CC substance inhibiting the formation of a complex between TAB1 and XIAP as
 CC well as between XIAP and TGF-beta (transforming growth factor-beta) type
 CC I and/or type II receptor is useful as a drug.

Query Match 100.0%; Score 359; DB 21; Length 504;
 Best Local Similarity 100.0%; Pred. No. 3.1e-37;
 Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QSTPLTQSTNTHTQSSSSDGLFRSPAHSLPPGDRGRVEPYVDFAEFRLWSVDHG 60
 DB 437 QSTPLTQSTNTHTQSSSSDGLFRSPAHSLPPGDRGRVEPYVDFAEFRLWSVDHG 496

QY 61 EQSVVTAP 68
 DB 497 EQSVVTAP 504

RESULT 6
 AAY09550
 ID AAY09550 standard; Protein; 513 AA.

XX AC AAY09550;
 XX 21-JUL-1999 (first entry)
 XX Human TAB1 protein SEQ ID NO:43.
 XX Human; TAB1; TAK1; screening; inhibition; TGF-beta;
 XX transforming growth factor beta.

XX OS Homo sapiens.
 XX WO9921010-A1.
 XX 29-APR-1999.
 XX 22-OCT-1998; 98WO-JP04796.
 XX 22-OCT-1997; 97JP-0290188.
 XX (CHUS) CHUGAI SEIYAKU KK.

XX Ontomo T, Ono K, Tsuchiya M;
 XX WPI; 1999-312645/26.
 XX N-PSDB; AAX56310.

PT Screening for TGF- beta inhibitory substances, which are useful as
 PT drugs for treatment of diseases relating to its disorder

XX Example 13; Page 186-188; 195pp; Japanese.

XX A method has been developed for screening for substances which inhibit
 CC the binding of TAK1 polypeptide to TAB1 polypeptide. The method
 CC comprises: (a) contacting the polypeptide in the presence of a sample;
 CC and (b) detecting the amount of bound polypeptide, in which the sample
 CC can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming
 CC growth factor (TGF)-beta inhibitory substances can be used in drugs for
 CC indications e.g. as TGF-beta signal transmission inhibitors or
 CC activators, or extracellular matrix protein production enhancement
 CC inhibitors, or activators, or cell proliferation prevention inhibitors or
 CC activators, or monocyte migration inhibitors or activators, or
 CC physiological activity induction inhibitors or activators, or
 CC immunosuppression inhibitors or activators, or amyloid beta protein

CC precipitation inhibitors or activators, and such substances can also be
 CC inhibitors of the TAK1 polypeptide function, particularly kinase
 CC activity. The present sequence represents human TAB1.

XX Sequence 513 AA;

Query Match 100.0%; Score 359; DB 20; Length 513;
 Best Local Similarity 100.0%; Pred. No. 3.2e-37;
 Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QSTPLTQSTNTHTQSSSSDGLFRSPAHSLPPGDRGRVEPYVDFAEFRLWSVDHG 60
 DB 446 QSTPLTQSTNTHTQSSSSDGLFRSPAHSLPPGDRGRVEPYVDFAEFRLWSVDHG 505

QY 61 EQSVVTAP 68
 DB 506 EQSVVTAP 513

RESULT 7
 AAY09546
 ID AAY09546 standard; Protein; 517 AA.

XX AC AAY09546;
 XX 21-JUL-1999 (first entry)
 XX Human TAB1-FLAG protein.
 XX Human; TAB1; TAK1; screening; inhibition; TGF-beta;
 XX transforming growth factor beta.

XX OS Homo sapiens.
 XX Synthetic.
 XX WO9921010-A1.
 XX 29-APR-1999.
 XX 22-OCT-1998; 98WO-JP04796.
 XX 22-OCT-1997; 97JP-0290188.

XX (CHUS) CHUGAI SEIYAKU KK.

XX Ontomo T, Ono K, Tsuchiya M;
 XX WPI; 1999-312645/26.
 XX N-PSDB; AAX56282.

PT Screening for TGF- beta inhibitory substances, which are useful as
 PT drugs for treatment of diseases relating to its disorder

XX Example 1; Page 163-166; 195pp; Japanese.

XX A method has been developed for screening for substances which inhibit
 CC the binding of TAK1 polypeptide to TAB1 polypeptide. The method
 CC comprises: (a) contacting the polypeptide in the presence of a sample;
 CC and (b) detecting the amount of bound polypeptide, in which the sample
 CC can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming
 CC growth factor (TGF)-beta inhibitory substances can be used in drugs for
 CC indications e.g. as TGF-beta signal transmission inhibitors or
 CC activators, or extracellular matrix protein production enhancement
 CC inhibitors, or activators, or cell proliferation prevention inhibitors or
 CC activators, or monocyte migration inhibitors or activators, or
 CC physiological activity induction inhibitors or activators, or
 CC immunosuppression inhibitors or activators, or amyloid beta protein
 CC precipitation inhibitors or activators, and such substances can also be
 CC inhibitors of the TAK1 polypeptide function, particularly kinase
 CC activity. The present sequence represents TAB1-FLAG from an example of
 CC the present invention.

XX Sequence 517 AA;

Query Match 100.0%; Score 359; DB 20; Length 517;
Best Local Similarity 100.0%; Pred. No. 3.2e-37;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QSPFTLTQSTNTHTQSSSSSDGGLFRSRPAHSLPPGSDGRVEPYVDPAEFYRLMSVDHG 60
Db 437 QSPFTLTQSTNTHTQSSSSSDGGLFRSRPAHSLPPGSDGRVEPYVDPAEFYRLMSVDHG 496
Qy 61 EOSVVTAP 68
Db 497 EOSVVTAP 504

RESULT 8

AAB56692 ID AAB56692 standard; Protein; 84 AA.

XX AC AAB56692;

XX DT 13-MAR-2001 (first entry)

XX DE Human prostate cancer antigen protein sequence SEQ ID NO:1270.

XX KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
XX KW neuroprotective; cytoskeletal; cardioprotective; immunomodulatory; muscular;
XX KW vulnary; gastrointestinal; nephrotoxic; antileukemic; gynaecological;
XX KW antibacterial; gene therapy; neutral; immune; reproductive; renal;
XX KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
XX KW wound; infectious disease.

XX OS Homo sapiens.

XX PN W020005174-A1.

XX PD 21-SEP-2000.

XX PF 08-MAR-2000; 2000WO-US05988.

XX PR 12-MAR-1999; 99US-0124270.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PA (ROSE/) ROSEN C A.

XX PI Rosen CA, Ruben SM;

XX WPI: 2000-587513/55.

XX DR N-PSDB; AAF15895.

XX PT Prostate cancer associated gene sequences, referred to as prostate
XX PT cancer antigens, useful for treatment, prevention, and diagnosis of
XX PT disorders such as prostate cancer -

XX PS Claim 11; Page 1693; 2338pp; English.

XX CC AAF15566 to AAF15505 encode the human prostate cancer associated
XX CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.

XX CC The prostate cancer antigens can have neuroprotective, cytoskeletal,
XX CC cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,
XX CC nephrotoxic, antileukemic, gynaecological and antibacterial activities,
XX CC and can be used in gene therapy. The prostate cancer antigen
XX CC polynucleotides may be used for detection of prostate cancer, chromosome
XX CC identification, as chromosome markers, and for numerous other diagnostic
XX CC or research purposes. The prostate cancer antigens may be used to treat
XX CC disorders such as neural, immune, muscular, reproductive,
XX CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
XX CC disorders, wounds, and infectious diseases. AAF15506 to AAF15514 to
XX CC AAB57303 represent sequences used in the exemplification of the present
XX CC invention.

XX SQ Sequence 84 AA;

Query Match 98.1%; Score 352; DB 21; Length 84;

Best Local Similarity 98.5%; Pred. No. 2.4e-37;
Matches 67; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QSPFTLTQSTNTHTQSSSSSDGGLFRSRPAHSLPPGSDGRVEPYVDPAEFYRLMSVDHG 60
Db 17 QSPFTLTQSTNTHTQSSSSSDGGLFRSRPAHSLPPGSDGRVEPYVDPAEFYRLMSVDHG 76
Qy 61 EOSVVTAP 68
Db 77 EOSVVTAP 84

RESULT 9

AA09549 ID AA09549 standard; peptide; 16 AA.

XX AC AA09549;

XX DT 21-JUL-1999 (first entry)

XX DE Human TAB1 peptide TAB1C-2.

XX KW Human; TAB1; screening; inhibition; TGF-beta;
XX KW transforming growth factor beta.

XX OS Homo sapiens.

XX PN W09921010-A1.

XX PD 29-APR-1999.

XX PF 22-OCT-1998; 98WO-JP04796.

XX PR 22-OCT-1997; 97JP-0290188.

XX PA (CHUS) CHUGAI SEIYAKU KK.

XX PI Ohtomo T, Ono K, Tsuchiya M;

XX WPI: 1999-312645/26.

XX PT Screening for TGF- beta inhibitory substances, which are useful as
XX PT drugs for treatment of diseases relating to its disorder

XX PS Example 13; Page 182; 195pp; Japanese.

XX CC A method has been developed for screening for substances which inhibit
XX CC the binding of TAK1 polypeptide to TAB1 polypeptide. The method
XX CC comprises: (a) contacting the polypeptide in the presence of a sample;
XX CC and (b) detecting the amount of bound polypeptide, in which the sample
XX CC can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming
XX CC growth factor (TGF)-beta inhibitory substances can be used in drugs for
XX CC indications e.g. as TGF-beta signal transduction inhibitors or
XX CC activators, or extracellular matrix protein production enhancement
XX CC inhibitors or activators, or cell proliferation prevention inhibitors or
XX CC activators, or monocyte migration inhibitors or activators, or
XX CC physiological activity induction inhibitors or activators, or
XX CC immunosuppression inhibitors or activators, or amyloid beta protein
XX CC precipitation inhibitors or activators, and such substances can also be
XX CC activators of the TAK1 polypeptide function, particularly kinase
XX CC activity. The present sequence represents a peptide from an example
XX CC of the present invention.

XX SQ Sequence 16 AA;

Query Match 21.4%; Score 77; DB 20; Length 16;

Best Local Similarity 100.0%; Pred. No. 0.0046;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QSPFTLTQSTNTHTQ 15
Db 2 QSPFTLTQSTNTHTQ 16

```
RESULT 10
ABB70838
ID ABB70838 standard; Protein; 1162 AA.
XX
XX ABB70838;
AC
XX
XX 26-MAR-2002 (first entry)
DT
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 39306.
DE
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
KW
KW pharmaceutical.
XX
XX Drosophila melanogaster.
OS
XX
XX WO200171042-A2.
PN
XX
XX 27-SEP-2001.
PD
XX
XX 23-MAR-2001; 2001WO-US09231.
PF
XX
XX 23-MAR-2000; 2000US-191637P.
PR
XX
XX 11-JUL-2000; 2000US-0614150.
PP
XX
XX (PEKE ) PE CORP NY.
PA
XX
XX Venter JC, Adams M, Li PWD, Myers BW;
PI
XX
XX WPI; 2001-656860/75.
DR
XX
XX N-PSDB; ABL14941.
DD
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
PT
XX
XX Disclosure; SEQ ID NO 39306; 21pp + Sequence Listing; English.
PS
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABBS7737-ABBS72072).
CC
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 1162 AA;
SQ
Query Match 19.5%; Score 70; DB 22; Length 1162;
Best Local Similarity 32.8%; Pred. No. 9.8;
Matches 22; Conservative 9; Mismatches 24; Indels 12; Gaps 2;
QY 3 PTLTLOSTWTHQSSSSSDGGLFRS-RAHSLPGEDGRVEPYVDFAEFYRLMSVDHGE 61
Db 960 PVLQPKSPSTTLTCHSSSSAGSAGYQYAPGMLPPR-----ADVARTLSLSNGS 1008
QY 62 QSVVTAP 68
Db 1009 SSEVTSP 1015
RESULT 11
AAG77804
ID AAG77804 standard; Protein; 199 AA.
XX
XX AAG77804;
AC
XX
XX 03-DEC-2001 (first entry)
DT
XX
```

DE Human HT4SG64 serine/threonine phosphatase protein sequence.

XX Human; HT4SG64; serine/threonine phosphatase; PSPase; vaccine; gene therapy; PSPase expression; PSPase expression; PSPase modulation; immune disorder; autoimmune disorder; Wiscott-Aldrich syndrome; Chediak-Higashi syndrome; Hashimoto's thyroiditis; multiple sclerosis; inflammation; Crohn's disease; inflammatory bowel disease; appendicitis; rheumatoid arthritis; cellular proliferative disorder; lymphoma; lung cancer; intestinal cancer; cardiovascular disorder; aneurysm; Scimitar syndrome; Ebstein's anomaly.

OS Homo sapiens.

XX

PH Key Location/Qualifiers

FT Region 70..84

FT /note= "Immunogenic region 1"

FT Region 114..121

FT /note= "Immunogenic region 2"

FT Region 126..136

FT /note= "Immunogenic region 3"

FT Region 159..164

FT /note= "Immunogenic region 4"

FT Region 186..199

FT /note= "Immunogenic region 5"

XX

PN WO200164703-A1.

XX

XX 07-SEP-2001.

PD

XX 28-FEB-2001; 2001WO-US06256.

PF

XX 02-MAR-2000; 2000US-0186350.

PR

XX (HUMA-) HUMAN GENOME SCI INC.

XX

PA Ebner R, Ruben SM;

XX

PI WPI; 2001-530113/58.

XX

DR N-PSDB; AAH78731.

DR

XX Nucleic acids encoding serine/threonine phosphatase polypeptides, useful for preventing, diagnosing and/or treating, e.g. Crohn's disease, lung cancer and Scimitar syndrome -

PT

PT

PT

XX

XX Claim 11; Page 323-324; 335pp; English.

PS

XX The present sequence represents the human HT4SG64 protein which is claimed in the invention. The invention comprises novel human serine/threonine phosphatase (PSPase) polypeptides and polynucleotides. The PSPase polynucleotides and polypeptides of the invention may be used in the prevention (vaccine), diagnosis and treatment (gene therapy) of diseases associated with inappropriate PSPase expression. The PSPase polynucleotides of the invention may be used as DNA probes to detect and quantitate the presence of similar nucleic acids in samples. The PSPase polypeptides may be used as antigens in the production of antibodies against the PSPase polypeptides and in assays to identify modulators of PSPase expression and activity. The anti-PSPase antibodies and antagonists may also be used to down regulate expression and activity, the anti-PSPase antibodies may also be used as diagnostic agents for detecting the presence of PSPase polypeptides in samples. Disorders that may be prevented, diagnosed and/or treated by the invention are: immune/autoimmune disorders (e.g. Wiscott-Aldrich syndrome, Chediak-Higashi syndrome, Hashimoto's thyroiditis and multiple sclerosis); inflammatory conditions (e.g. Crohn's disease, inflammatory bowel disease, appendicitis and rheumatoid arthritis); cellular proliferative disorders (e.g. lymphoma, lung and intestinal cancers); and cardiovascular disorders (e.g. Scimitar syndrome, Ebstein's anomaly and aneurysm).

XX

SQ Sequence 199 AA;

Query Match 19.4%; Score 69.5; DB 22; Length 199;

Best Local Similarity 41.9%; Pred. No. 1.1;

Matches 18; Conservative 3; Mismatches 17; Indels 5; Gaps 1;
QY 1 QSPFLTOSTNTHTOSSSSSDGGLFRSRPAHSIPGEGRVPEYVDFAE 38
| | | | | : : : : :
| | | | | : : : : :
Db 119 QSPFLTIDQTRHRSRDAFSSISGCSKFTAVKRMADKLPIVQ 161

RESULT 12
AAU47075
ID AAU47075 standard; Protein; 113 AA.
XX
AC AAU47075;
DT 27-FEB-2002 (first entry)
DE Propionibacterium acnes immunogenic protein #7971.
XX

KM SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KM inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KM dermatological; osteopathic; neuroprotectant.

OS Propionibacterium acnes.

XX WO200181581-A2.

XX PD 01-NOV-2001.

XX PF 20-APR-2001; 2001WO-US12865.

XX PR 21-APR-2000; 2000US-199047P.

XX PR 02-JUN-2000; 2000US-208841P.

XX PR 07-JUL-2000; 2000US-216747P.

PA (CORI-) CORIXA CORP.

PI Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhacia A;

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

DR WPI; 2001-616774/71.

DR N-PSDB; AASS5936.

XX Example 1; SEQ ID No 8270; 1069pp; English.

CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the inflammatory
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 113 AA;
Query Match 19.1%; Score 68.5; DB 22; Length 113;
Best Local Similarity 35.4%; Pred. No. 0.72;

Matches 17; Conservative 8; Mismatches 22; Indels 1; Gaps 1;
QY 3 PTLTOSTNTHTOSSSSSDGGLFRSRPAHSIPGEGRVPEYVDFAE 50
| | | | | : : : : :
| | | | | : : : : :
Db 61 PTLTAATSGQMMNTBPGT-GGLVGPSPMKSTRPRPHDHGIMTYPRFAD 107

RESULT 13
ABG16477
ID ABG16477 standard; Protein; 739 AA.
XX
AC ABG16477;
DT 18-FEB-2002 (first entry)
DE Novel human diagnostic protein #16468.
XX

KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.

XX WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI; 2001-639362/73.

DR N-PSDB; AAS80664.

XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity -

XX Claim 20; SEQ ID No 46836; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 739 AA;
Query Match 18.8%; Score 67.5; DB 22; Length 739;
Best Local Similarity 31.1%; Pred. No. 11;
Matches 23; Conservative 12; Mismatches 20; Indels 19; Gaps 3;

```
QY      2  SPTLTLOSTNTHTQSSSSSDG-----CLFRSRPAH-----SLPPGDGRVPEYVDF 48
Db      119  SSTGSLPPTNTNTSEGATSLIILPTISGSGRPLNPVTQAPLPFGWEQRVD----- 173
QY      49  AEFYRLWSVDHGEQ 62
Db      174  -QHGRVYVDHVEK 186

RESULT 14
AAG68173
ID  AAG68173 standard; Protein; 739 AA.
XX
AC  AAG68173;
XX
DT  25-JAN-2002 (first entry)
XX
DE  Atrophin-1 interacting protein (AIP4) SEQ ID NO:89.
XX
KW  Human; high bone mass; HBM gene; Zmax1 gene; chromosome 11; 11q13.3;
KW  sequence tagged site; STS; osteoporosis; osteopathic; gene therapy;
KW  antisenese therapy; vaccine; bone disorder; Paget's disease;
KW  sclerostosis; osteomalacia; fibrous dysplasia.
XX
OS  Homo sapiens.
XX
PN  WO200177327-A1.
XX
PD  18-OCT-2001.
XX
PF  21-JUN-2000; 2000WO-US16951.
XX
PR  05-APR-2000; 2000US-0543771.
PR  05-APR-2000; 2000US-0544398.
XX
PA  (GENO-) GENOME THERAPEUTICS CORP.
XX
PI  Carulli JP, Little RD, Recker RR, Johnson ML;
XX
DR  WPI; 2001-657171/75.
XX
PT  New high bone mass (HBM) and Zmax1 genes and proteins useful for
PT  modulating bone mass for the treatment of e.g. osteoporosis -
XX
PS  Claim 76; Page 392-394; 443pp; English.
XX
CC  The present invention describes the human Zmax1 gene and the high bone
CC  mass (HBM) gene, which are found on chromosome 11q13.3. The Zmax1 and
CC  HBM genes have osteopathic activities. The genes can be used in gene
CC  therapy, antisenese therapy and in the production of vaccines. They
CC  can be used in the diagnosis and treatment of bone disorders including
CC  osteoporosis, Paget's disease, sclerostosis, osteomalacia and fibrous
CC  dysplasia. ABA82038 to ABA82700 and AAG68168 to AAG68193 represent
CC  sequences used in the exemplification of the present invention.
XX
SQ  Sequence 739 AA;

Query Match      18.8%; Score 67.5; DB 22; Length 739;
Best Local Similarity 31.1%; Pred. No. 11;
Matches 23; Conservative 12; Mismatches 20; Indels 19; Gaps 3;

QY      2  SPTLTLOSTNTHTQSSSSSDG-----CLFRSRPAH-----SLPPGDGRVPEYVDF 48
Db      119  SSTGSLPPTNTNTSEGATSLIILPTISGSGRPLNPVTQAPLPFGWEQRVD----- 173
QY      49  AEFYRLWSVDHGEQ 62
Db      174  -QHGRVYVDHVEK 186

RESULT 15
AAY30948
ID  AAY30948 standard; Protein; 852 AA.
```

```
XX
AC  AAY30948;
XX
DT  21-OCT-1999 (first entry)
XX
DE  Human E3 ubiquitin protein ligase protein.
XX
KW  E3 ubiquitin protein ligase; h-E3 UPL; antiinflammatory; antidiabetic;
KW  immunosuppressive; neuroprotective; cytosolic; antiarthritic; cardiac;
KW  immunomodulator; antiviral; treatment; screening; gene therapy; cancer;
KW  inflammation; autoimmune disease; neurological disease; apoptosis;
KW  endothelial cell; proliferation; differentiation; angiogenesis; cachexia;
KW  peripheral vascular disease; hematopoietic disorder; arthritis; leukemia;
KW  pulmonary disorder; diabetes; viral infection; human.
XX
OS  Homo sapiens.
XX
FH  Key Location/Qualifiers
FT  Domain 275..306
FT  /label= WWI
FT  /note= "WW protein interaction domain I"
FT  Domain 307..340
FT  /label= WWII
FT  /note= "WW protein interaction domain II"
FT  Domain 386..420
FT  /label= WWIII
FT  /note= "WW protein interaction domain III"
FT  Domain 427..460
FT  /label= WWIV
FT  /note= "WW protein interaction domain IV"
XX
PN  WO9940201-A1.
XX
PD  12-AUG-1999.
XX
PF  02-FEB-1999; 99WO-GB00353.
XX
PR  30-APR-1998; 98US-0070060.
PR  05-FEB-1998; 98US-0073839.
XX
PA  (ZENE ) ZENECA LTD.
XX
PI  Chidyal N, Hustad CM;
XX
DR  WPI; 1999-508506/42.
DR  N-PSDB; AAZ09235, AAY09236.
XX
PT  New human proteolytic accessory enzyme and its modulators useful
PT  for treating disease conditions like inflammation or autoimmune
PT  diseases
XX
PS  Claim 1; Fig 3; 95pp; English.
XX
CC  This invention describes a novel human polynucleotide (I) which encodes a
CC  E3 ubiquitin protein ligase, h-E3 UPL (I1). The products of the invention
CC  have antiinflammatory, immunosuppressive, neuroprotective, cytosolic,
CC  antiarthritic, immunomodulator, antidiabetic, antiviral and cardiac
CC  activity. The products of the invention can also be used for treating
CC  patients with disorders mediated by the biological and/or pharmacological
CC  activity of h-E3 UPL. The nucleic acid sequences encoding E3 UPL are used
CC  in expression systems as assay for agonists and antagonists for the E3
CC  UPL protein. The E3 UPL protein is used in screening assays to identify
CC  blockers and antagonists. They are also used in gene therapy. Specific
CC  modulation of biological and/or pharmacological activity of novel h-E3
CC  UPL via administration of a modulator or heterologous expression, is used
CC  for treating physiological conditions like inflammation, autoimmune
CC  diseases, neurological disease, apoptosis, endothelial cell physiology
CC  (e.g., proliferation, differentiation), peripheral vascular disease,
CC  angiogenesis, cancer, hematopoietic disorders, arthritis, cachexia,
CC  leukemia, pulmonary disorders, diabetes and viral infection. The
CC  nucleotide sequences which encode h-E3 UPL may also be employed in
CC  analysis to map chromosomal location e.g., screening for functional
CC  association with disease markers. They are also used as screening tools
```


CC in the identification of appropriate human subjects and patients for
 CC therapeutic clinical trials. The sequences can also be used to detect the
 CC presence of the mRNA transcripts in a patient or to monitor the
 CC modulation of transcripts during treatment. This sequence represents the
 CC human E3 ubiquitin protein ligase protein described in the method of the
 CC invention.
 CC
 XX

SQ Sequence 852 AA;

Query Match 18.8%; Score 67.5; DB 20; Length 852;

Best Local Similarity 31.1%; Pred. No. 14; Mismatches 20; Indels 19; Gaps 3;

Matches 23; Conservative 12; Mismatches 20; Indels 19; Gaps 3;

QY 2 SPTLTQSTNTHTQSSSSSDG-----GLFRSRPAH-----SLPPGEDGRVEPYVDF 48
 Db 232 SSTGSLPTNTNTNTSGATGGLIPLTISGSGPRPLNPVTQAFLPPGWGEQRVD----- 286

QY 49 AEFYRLMSVDHGEQ 62

Db 287 -QHGRVYVVDHVEK 299

Search completed: December 10, 2002, 07:30:42
 Job time : 87 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 10, 2002, 07:27:32 ; Search time 52 seconds

(without alignment)
125.714 Million cell updates/sec

Title: US-09-830-144-4_COPY_437_504

Sequence: 1 QSPRLTQSTHTTQSSSSS.....AEFRLMSVDHGEQSVTAP 68

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1.*
2: p1r2.*
3: p1r3.*
4: p1r4.*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	63.5	17.7	243	D84792	probable protein t
2	62	17.3	541	S51799	nucleoporin NUP57
3	62	17.3	679	H95036	glycosyl hydrolase
4	62	17.3	737	D97907	alpha-xylosidase (
5	62	17.3	4957	T03455	ALR protein - huma
6	62	17.3	5262	T03454	ALR protein - huma
7	61.5	17.1	238	T52505	hypothetical prote
8	61.5	17.1	737	T52505	hypothetical prote
9	61	17.0	267	S74415	S-protein secretio
10	61	17.0	665	E82506	hypothetical prote
11	61	17.0	909	T06246	proteinase II YCA0
12	61	17.0	951	E88042	aspartate kinase (
13	61	17.0	1511	A53151	protein P56D12.6 (
14	60	16.7	1260	S60896	pleiotropic drug r
15	60	16.7	1576	AE0249	agglutinin-like pr
16	59.5	16.6	1369	T32338	probable hemolysin
17	59	16.4	173	G86861	hypothetical prote
18	59	16.4	329	S61884	hypothetical prote
19	59	16.4	330	S61883	carbonate dehydrat
20	59	16.4	330	S61882	carbonate dehydrat
21	59	16.4	429	T38146	dihydrofolate redu
22	59	16.4	651	UC7705	death receptor-6 -
23	58.5	16.3	378	H72679	hypothetical prote
24	58.5	16.3	378	AA8161	hypothetical prote
25	58.5	16.3	608	T53269	prolactin receptor
26	58.5	16.3	1093	T38533	AF17 protein - hum
27	58.5	16.3	1099	T18257	phospholipase C -
28	58.5	16.3	2897	B48666	cell proliferation
29	58.5	16.3	3256	A48666	cell proliferation

30	58	16.2	142	2	D69891	ynek protein - Bac
31	58	16.2	277	2	D69158	sensory transducti
32	58	16.2	532	2	B35621	spore germination
33	57.5	16.0	172	2	T36107	probable serine/ar
34	57.5	16.0	232	1	A25108	homeotic protein H
35	57.5	16.0	366	2	A86392	hypothetical prote
36	57.5	16.0	394	2	H65010	hypothetical prote
37	57.5	16.0	396	2	A57090	CSA protein - huma
38	57.5	16.0	629	2	B83107	chemotactic transd
39	57.5	16.0	666	2	UC5708	villin-like protei
40	57	15.9	762	2	S67765	probable membrane
41	57	15.9	778	2	H84678	hypothetical prote
42	56.5	15.7	276	2	T06331	photosystem II 22k
43	56.5	15.7	283	2	H83860	patocarbonate synth
44	56.5	15.7	343	2	E83673	soyabitol dehydroge
45	56.5	15.7	457	2	B64790	yc1B protein - Esc

ALIGNMENTS

RESULT 1

D84792
probable protein translocase [imported] - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C/Accession: D84792
R/Lin: X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; I

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vankken, S.E.; Unayam, L.; Tallon, L.

gens, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J

Nature 402, 761-768, 1999

A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A/Reference number: A84420; MUID:20083487; PMID:10617197

A/Accession: D84792

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-243 <STO>

A/Cross-references: GB:AE002093; NID:94056494; PIDN:AC98060.1; GSPDB:GN00139

C/Genetics:

A/Map position: 2

A/Map position: 2

A/Map position: 2

A/Map position: 2

A/Map position: 2

A/Map position: 2

A/Map position: 2

A/Map position: 2

A/Map position: 2

A/Map position: 2

A/Map position: 2

A/Map position: 2

A/Map position: 2

A/Map position: 2

A/Map position: 2

Query Match 17.7%; Score: 63.5; DB 2; Length 243;
Best Local Similarity 39.0%; Pred. No. 7.9;
Matches 16; Conservative 6; Mismatches 18; Indels 1; Gaps 1;

QY 8 QSTNTHQSSSSSDGGLFRSPAHSLPCEGDRVPEYDF 48

Db 193 QNONTASSSSSSWFGGLF-DKKKEVQPSSEKTEVLESF 232

RESULT 2

S51799

nucleoporin NUP57 - yeast (Saccharomyces cerevisiae)

N/Alternate names: protein G6320; protein YGR119C

C/Species: Saccharomyces cerevisiae

C/Date: 27-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 21-Jul-2000

C/Accession: S51799; S51800; S64428; S64427; S55976; S72192

R/Schlauch, N.L. EMBL Data Library, August 1994

submitted to the EMBL Data Library, August 1994

A/Reference number: S51799

A/Accession: S51799

A/Molecule type: DNA

A/Residues: 1-541 <SCD>

A/Cross-references: EMBL:X8155; NID:9671635; PID:9671636

R/Grant: P.; Schlauch, N.; Tekotte, H.; Hurt, E.C.

EMBO J. 14, 76-87, 1995

A/Title: Functional interaction of NUP57 with a core nucleoporin complex consisting of

A/Reference number: S51800; MUID:9512954; PMID:7828598

A/Accession: S51800

A/Status: nucleic acid sequence not shown

A/Molecule type: DNA

A/Residues: 1-446; 'RL', 449-541 <GRA>

A/Residues: 1-446; 'RL', 449-541 <GRA>

A/Residues: 1-446; 'RL', 449-541 <GRA>

A/Residues: 1-446; 'RL', 449-541 <GRA>

A/Residues: 1-446; 'RL', 449-541 <GRA>

A/Residues: 1-446; 'RL', 449-541 <GRA>

A/Residues: 1-446; 'RL', 449-541 <GRA>

A/Residues: 1-446; 'RL', 449-541 <GRA>

A;Cross-references: EMBL:X81155
 R;Van Dyck, L.; Skala, J.; de Wergifosse, P.; Purnelle, B.; Talla, E.; Nawrocki, A.; Del
 submitted to the Protein Sequence Database, May 1996
 A;Reference number: S64428
 A;Accession: S64428
 A;Molecule type: DNA
 A;Residues: 1-353 <VAN>
 A;Cross-references: EMBL:272904; MIPS:YGR119C
 A;Experimental source: strain S288C
 R;Hansen, M.; Albers, M.; Backes, U.; Coblentz, A.; Leuther, H.; Neu, R.; Schreier, A.; Sch
 submitted to the Protein Sequence Database, May 1996
 A;Reference number: S64417
 A;Accession: S64427
 A;Molecule type: DNA
 A;Residues: 300-541 <HAN>
 A;Cross-references: EMBL:272904; MIPS:YGR119C
 A;Experimental source: strain S288C
 R;van Dyck, L.; Goffeau, A.
 submitted to the EMBL Data Library, December 1994
 A;Description: Genes for an asn synthase, a GLFG-motif nucleoporin and a putative homeob
 e new ORFs, remnants of Ty and three tRNA genes.
 A;Reference number: S55976
 A;Accession: S55976
 A;Molecule type: DNA
 A;Residues: 1-353 <VAV>
 A;Cross-references: EMBL:X83099
 R;Hansen, M.; Albers, M.; Backes, U.; Coblentz, A.; Leuther, H.; Neu, R.; Schreier, A.; Sch
 Yeast 12, 1273-1277, 1996
 A;Title: The sequence of a 23.4 kb segment on the right arm of chromosome VII from Sacch
 A;Reference number: S72179; MUID:97061913; PMID:8905931
 A;Accession: S72192
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-541 <HAW>
 A;Cross-references: EMBL:272904; NID:G1323192; PIDN:CAA97129.1; PID:G1323193
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, May 1996
 C;Genetics:
 A;Gene: SGD:NUP57
 A;Cross-references: SGD:S0003351; MIPS:YGR119C
 A;Map position: 7R

Query Match 17.3%; Score 62; DB 2; Length 541;
 Best Local Similarity 50.0%; Pred. No. 30;
 Matches 12; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 8 QSTNTHTQSSSSSDGGLFRSRPA 31
 Db 60 QATNTGSGNQSSGTGGGLFCNRP 83

RESULT 3
 H95036
 glycosyl hydrolase, family 31 SP0312 [imported] - Streptococcus pneumoniae (strain TIGR4
 C;Species: Streptococcus pneumoniae
 C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
 C;Accession: H95036
 R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
 on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.K.; Radune, D.; Holtzapple,
 nson, T.; Hickey, E.K.; Holt, I.E.
 Science 293, 498-506, 2001
 A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
 A;Reference number: A95000; MUID:21357209; PMID:11463916
 A;Accession: H95036
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-679 <KOR>
 A;Cross-references: GB:AE005672; PIDN:AAK74489.1; PID:G14971785; GSPDB:GN00164; TIGR.SP4
 A;Experimental source: strain TIGR4
 C;Genetics:
 A;Gene: SP0312

Query Match 17.3%; Score 62; DB 2; Length 679;

Best Local Similarity 32.8%; Pred. No. 39;
 Matches 21; Conservative 4; Mismatches 23; Indels 16; Gaps 3;

QY 13 HTQSSSSSDGGLFRSRPA---HSLPGEQG-----RVEPYVDFAEF---YRLWS 56
 Db 289 HYQDSCKNAEGGLILSRVAGPGSHRYPVGSGDTIISNLSLRFQPYFTATASNIGYSWWS 348

QY 57 VDHG 60
 Db 349 HDIG 352

RESULT 4
 D97907
 alpha-xylosidase (EC 3.2.1.1-) [imported] - Streptococcus pneumoniae (strain R6)
 C;Species: Streptococcus pneumoniae
 C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
 C;Accession: D97907
 R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
 e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
 Y, P.; Sun, P.M.; Winkler, M.E.
 J. Bacteriol. 183, 5709-5717, 2001
 A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaaskunas, S.R.;
 A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
 A;Reference number: A97872; MUID:21429245; PMID:11544234
 A;Accession: D97907
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-737 <KUR>
 A;Cross-references: GB:AE007317; PIDN:AAK99088.1; PID:G15457836; GSPDB:GN00174
 C;Genetics:
 A;Gene: xylS
 C;Keywords: glycosidase; hydrolase

Query Match 17.3%; Score 62; DB 2; Length 737;
 Best Local Similarity 32.8%; Pred. No. 42;
 Matches 21; Conservative 4; Mismatches 23; Indels 16; Gaps 3;

QY 13 HTQSSSSSDGGLFRSRPA---HSLPGEQG-----RVEPYVDFAEF---YRLWS 56
 Db 347 HYQDSCKNAEGGLILSRVAGPGSHRYPVGSGDTIISNLSLRFQPYFTATASNIGYSWWS 406

QY 57 VDHG 60
 Db 407 HDIG 410

RESULT 5
 T03455
 ALR protein - human
 C;Species: Homo sapiens (man)
 C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 21-Jul-2000
 C;Accession: T03455
 R;Prasad, R.; Zhadanov, A.B.; Sedkov, Y.; Bullrich, F.; Druck, T.; Rallapalli, R.; Yano,
 Oncogene 15, 549-560, 1997
 A;Title: Structure and expression pattern of human ALR, a novel gene with strong homology
 A;Reference number: Z14954; MUID:97388474; PMID:9247308
 A;Accession: T03455
 A;Status: preliminary; translated from GB/EMBL/DDJ
 A;Molecule type: mRNA
 A;Residues: 1-4957 <PRA>
 A;Cross-references: EMBL:AF010404; NID:G2358286; PIDN:AAK51735.1; PID:G2358287
 C;Genetics:
 A;Gene: ALR
 A;Map position: 12
 A;Superfamily: human ALR protein
 C;Keywords: alternative splicing

Query Match 17.3%; Score 62; DB 2; Length 4957;
 Best Local Similarity 33.3%; Pred. No. 3.9e+02;
 Matches 26; Conservative 8; Mismatches 22; Indels 22; Gaps 6;

QY 1 QSTTLTLQSTNTHQTQSS---SSSSDG---GLFRSRPAHSL-PPGEDGRVEPYVDFAEFY 52

Db 1996 QSTNYTATGNFHPSSGSLPGSSSTGSESYGLSPLRPSVLPFPAPDGL-PY----- 2047
 QY 53 RLMSVDHG--EQSVVTP 68
 Db 2048 ----LSHGASQSGITSP 2061

RESULT 6

T03454
 ALR protein - human
 C/Species: Homo sapiens (man)
 C/Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 21-Jul-2000
 C/Accession: T03454
 R/Prasad, R.; Zhadanov, A.B.; Sedkov, Y.; Bullrich, F.; Druck, T.; Rallapalli, R.; Yano, Oncogene 15, 549-560, 1997
 A/Title: Structure and expression pattern of human ALR, a novel gene with strong homolog
 A/Reference number: Z14954; PMID:97388474; PMID:9247308
 A/Accession: T03454
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-5262 <PRA>
 A/Cross-references: EMBL:AF010403; NID:g2358284; PIDN:AAC51734.1; PID:g2358285
 C/Genetics:
 A/Gene: ALR
 A/Map position: 12
 C/Superfamily: human ALR protein
 C/Keywords: alternative splicing

Query Match 17.3%; Score 62; DB 2; Length 5262;
 Best Local Similarity 33.3%; Pred. No. 41e+02;
 Matches 26; Conservative 8; Mismatches 22; Indels 22; Gaps 6;
 QY 1 QSPRLTQSTNTHTQSS-----SSSDG---GLFRSRPAHSL-PPGEDGRVEPVYDFAEFY 52
 Db 2301 QSTNYTATGNFHPSSGSLPGSSSTGSESYGLSPLRPSVLPFPAPDGL-PY----- 2352
 QY 53 RLMSVDHG--EQSVVTP 68
 Db 2353 ----LSHGASQSGITSP 2366

RESULT 7

T52505
 hypothetical protein B2023.10 [imported] - Neurospora crassa
 C/Species: Neurospora crassa
 C/Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
 C/Accession: T52505
 R/Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura, submitted to the Protein Sequence Database, September 2000
 A/Reference number: Z26053
 A/Accession: T52505
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-238 <SCH>
 A/Cross-references: EMBL:AF442164; GSPDB:GN00116; NCSP:B2023.10
 A/Experimental source: BAC clone B2023; strain OR74A
 C/Genetics:
 A/Gene: NCSP:B2023.10
 A/Map position: 6
 A/Introns: 188/2

Query Match 17.1%; Score 61.5; DB 2; Length 238;
 Best Local Similarity 31.6%; Pred. No. 13;
 Matches 24; Conservative 8; Mismatches 27; Indels 17; Gaps 3;
 QY 2 SPTTLTQSTNTHTQSSSSSDGGLFPR-----SRPAHSLPPGEDG-----RVEPVYDF 48
 Db 24 SPT-----ATNTSSSSSSSSSSSAFGVLPRLPISYPTALIGPRRGGPKTFCNKNGPVLDQ 79
 QY 49 AEFRLMSVDHGEQSV 64
 Db 80 TSAGSPYSVDCKQISI 95

RESULT 8

139547
 S-protein secretion D - Aeromonas hydrophila
 C/Species: Aeromonas hydrophila
 C/Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 08-Oct-1999
 C/Accession: I39547
 R/Thomas, S.R.; Trust, T.J.
 J. Bacteriol. 177, 3932-3939, 1995
 A/Title: A specific Pseudomonas sp. is required for the secretion of paracrystalline surface
 A/Reference number: A57354; PMID:95332195; PMID:7608063
 A/Accession: I39547
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-737 <RES>
 A/Cross-references: GB:I41682; NID:g950300; PIDN:AAA79322.1; PID:g1019923
 C/Genetics:
 A/Gene: spsd

Query Match 17.1%; Score 61.5; DB 2; Length 737;
 Best Local Similarity 37.8%; Pred. No. 48;
 Matches 17; Conservative 3; Mismatches 20; Indels 5; Gaps 1;
 QY 3 PTLTQSTNTHTQSSSSSDGGLFRRPRAHS-----LPPGEDGRV 42
 Db 385 PATGVNSTNTVTATGEGSGAGSLTAAAPATAPTSTAPRGEQGSV 429

RESULT 9

S74415
 hypothetical protein s110688 - Synechocystis sp. (strain PCC 6803)
 C/Species: Synechocystis sp.
 A/Variety: PCC 6803
 C/Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
 C/Accession: S74415
 R/Kaneko, T.; Sato, S.; Kori, H.; Tanaka, A.; Asanizu, E.; Nakamura, Y.; Miyajima, N.; O.K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda DNA Res. 3, 109-136, 1996
 A/Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp.
 A/Reference number: S74322; PMID:97061201; PMID:8905231
 A/Accession: S74415
 A/Status: nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-267 <KAN>
 A/Cross-references: EMBL:D64001; GB:AB001339; NID:g1001102; PIDN:BA10333.1; PID:g101098
 A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 17.0%; Score 61; DB 2; Length 267;
 Best Local Similarity 29.9%; Pred. No. 17;
 Matches 20; Conservative 12; Mismatches 21; Indels 14; Gaps 3;
 QY 1 QSP--LTLQSTNTHTQSSSSSDGGLFRRSR--PAHSLPPGEDGRVEPVYDFAEFYRLM 55
 Db 153 QEPNKNFTLRLATQAKQATGSLVGAESKSLVLRDVPPIKQKV-----YQLM 203
 QY 56 SYDHGEQ 62
 Db 204 AVIDGKK 210

RESULT 10

B82506
 proteinase II VCA0063 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
 C/Species: Vibrio cholerae
 C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
 C/Accession: B82506
 R/Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.; Chardon, D.; Ermolaeva, M.D.; Vamathevan, J.; Bess, S.; Qin, H.; Dragoi, I.; Sellers, P. Nature 406, 477-483, 2000
 A/Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

F:560-576/Domain: transmembrane #status predicted <TM2>
 F:612-628/Domain: transmembrane #status predicted <TM3>
 F:613-649/Domain: transmembrane #status predicted <TM4>
 F:663-679/Domain: transmembrane #status predicted <TM5>
 F:777-794/Domain: transmembrane #status predicted <TM6>
 F:888-1087/Domain: transmembrane #status predicted <TM7>
 F:905-912/Region: nucleotide-binding motif A (P-loop)
 F:1036-1252/Domain: transmembrane #status predicted <TM7>
 F:1270-1295/Domain: transmembrane #status predicted <TM8>
 F:1330-1346/Domain: transmembrane #status predicted <TM9>
 F:1358-1374/Domain: transmembrane #status predicted <TM10>
 F:1389-1405/Domain: transmembrane #status predicted <TM11>
 F:1478-1495/Domain: transmembrane #status predicted <TM12>

Query Match 17.0%; Score 61; DB 2; Length 1511;
 Best Local Similarity 38.8%; Pred. No. 1.3e+02;
 Matches 19; Conservative 6; Mismatches 16; Indels 8; Gaps 3;

QY 4 TLTLQSTNTHQTSSSSSDGGLFRSP---RPAHSLP--PGEDGRVDP 44
 DB 49 TLTAQSMQNSTQSAAPKSDAQSIFFSGVGVNPIFSDPEAPGYDPKLDP 97

RESULT 14

S60896
 agglutinin-like protein - yeast (Candida albicans)
 C/Species: Candida albicans
 C/Date: 27-Apr-1996 #sequence_revision 13-Mar-1997 #text_change 17-Mar-2000
 C/Accession: S60896
 R/Hoyer, L.L.; Scherer, S.; Shatzman, A.R.; Livi, G.P.
 Mol. Microbiol. 15, 39-54, 1995
 A/Title: Candida albicans ALS1: domains related to a Saccharomyces cerevisiae sexual ag
 A/Reference number: S60896; MUID:95272392; PMID:7752895
 A/Accession: S60896
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-1260 <HOY>
 A/Cross-references: EMBL:L25902; NID:G704426; PIDN:AAC41649.1; PID:G704427
 C/Superfamily: Yeast glucan 1,4-alpha-glucosidase homolog; glucan 1,4-alpha-glucosidase

Query Match 16.7%; Score 60; DB 2; Length 1260;
 Best Local Similarity 24.2%; Pred. No. 1.3e+02;
 Matches 23; Conservative 9; Mismatches 35; Indels 28; Gaps 3;

QY 2 SPTLTL-----QSTNTHQTSSSSSDGGLFRSPAHSL-----PP 36
 DB 651 NPTVTTEYWSQSYATTTTITAPPEGTDVTLIREPPNHTVTTEYWSQSYATTTVTAP 710
 QY 37 GEDGRV---EPYVDFAEFYRLMSVDHGQSVVTAP 68
 DB 711 GETDVTLIREPPNHTVTTEYWSQSYATTTVTAP 745

RESULT 15

AE0249
 probable hemolysin YPO2045 [imported] - Yersinia pestis (strain CO92)
 C/Species: Yersinia pestis
 C/Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
 C/Accession: AE0249
 R/Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
 demo-Tarrage, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
 11, M.; Rutherford, K.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
 Nature 413, 523-527, 2001
 A/Title: Genome sequence of Yersinia pestis, the causative agent of plague.
 A/Reference number: AB0001; MUID:21470413; PMID:11586360

A/Accession: AE0249
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-1576 <XUR>
 A/Cross-references: GB:AL590842; PIDN:CAC90857.1; PID:G15980056; GSPDB:GN00175
 C/Genetics:
 A/Gene: YPO2045

Query Match 16.7%; Score 60; DB 2; Length 1576;
 Best Local Similarity 25.4%; Pred. No. 1.7e+02;
 Matches 17; Conservative 10; Mismatches 20; Indels 20; Gaps 3;

QY 5 LTLQSTNTHQTSSSSSDGGLFRSPAHSLPPEGDRVPEYVDF---EF---YRLW 55
 DB 506 IDIQKNTHIHGAKIASGGL-----SIDAKGVYIIGASMLTSEFKDIDYNNW 554
 QY 56 SYDHRGEQ 62
 DB 555 GGAHGE 561

Search completed: December 10, 2002, 07:34:35
 Job time : 56 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 10, 2002, 06:49:56 ; Search time 42 Seconds
(without alignments)
67,152 Million cell updates/sec

Title: US-09-830-144-4_COPY_437_504

Perfect score: 359
Sequence: 1 QSPYTLTQSTNTHTQSSSSS.....AEFYRLMSVDHGEGSVTAP 68

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	359	100.0	504	1 TAB1_HUMAN	Q15750 homo sapien
2	63.5	17.7	243	1 IM17_ASATH	O99P35 arabidopsis
3	63.5	17.7	824	1 AD08_HUMAN	P78325 homo sapien
4	62	17.3	541	1 NU07_YEAST	P48837 saccharomyc
5	61	17.0	1119	1 AL03_CANAL	O74623 candida alb
6	61	17.0	1511	1 PDR5_YEAST	P33302 saccharomyc
7	60	16.7	226	1 OAZ_MOUSE	P54369 mus musculu
8	60	16.7	1260	1 AL01_CANAL	P46590 candida alb
9	59	16.4	328	1 CAHX_PLAAR	P46281 flaveria pr
10	59	16.4	330	1 CAHX_PLALI	P46512 flaveria li
11	59	16.4	330	1 CAHX_PLAAR	P46511 flaveria br
12	59	16.4	399	1 BR03_SHEEP	O97967 ovis aries
13	58.5	16.3	608	1 PRLR_MOUSE	O08501 mus musculu
14	58.5	16.3	1093	1 AF17_HUMAN	P55198 homo sapien
15	58.5	16.3	1099	1 PL01_CANAL	O13433 candida alb
16	58.5	16.3	3256	1 K167_HUMAN	P46013 homo sapien
17	58	16.2	142	1 YNEK_BACSU	P45711 bacillus su
18	58	16.2	276	1 CAPE_DROME	P46603 drosophila
19	58	16.2	532	1 SP07_DICDI	P26698 dictyostel
20	57.5	16.0	232	1 HXB4_XENLA	P09070 xenopus lae
21	57.5	16.0	396	1 CSA_HUMAN	O13216 homo sapien
22	57.5	16.0	686	1 V1LE_HUMAN	O13195 homo sapien
23	57	15.9	937	1 NU08_HUMAN	P52948 homo sapien
24	57	15.9	1377	1 NE01_RAT	P97603 rattus norv
25	57	15.9	1461	1 NE01_HUMAN	O92859 homo sapien
26	56.5	15.7	276	1 P8B3_LYCES	P54773 lycopersico
27	56.5	15.7	283	1 PANC_BACHD	O94086 bacillus ha
28	56.5	15.7	343	1 DH00_BACHD	O92901 bacillus ha
29	56.5	15.7	457	1 CUSC_ECOLI	P77211 escherichia
30	56.5	15.7	459	1 MCE1_YEAST	O04159 saccharomyc
31	56.5	15.7	463	1 PL0B_CARTI	O04213 carthamus t
32	56.5	15.7	922	1 NR01_RAT	O90019 rattus norv
33	56.5	15.7	923	1 NR01_HUMAN	O14786 homo sapien

34	56.5	15.7	1858	1 P3K2_DICDI	P54674 dictyostel
35	56	15.6	297	1 PR7A_SPRGR	P00776 streptomyce
36	56	15.6	315	1 SPY2_HUMAN	O43597 homo sapien
37	56	15.6	325	1 GBLP_SOYBN	O39836 glycine max
38	56	15.6	405	1 Y574_HUMAN	O60320 homo sapien
39	56	15.6	427	1 TR16_HUMAN	P08138 homo sapien
40	56	15.6	509	1 RUNT_DROME	P22814 drosophila
41	56	15.6	639	1 BARI_SCHCO	O92275 schizophy11
42	56	15.6	652	1 FX01_MOUSE	O91100 mus musculu
43	56	15.6	743	1 TLE2_HUMAN	O04725 homo sapien
44	56	15.6	776	1 AD28_MACFA	O9x516 macaca fasc
45	56	15.6	889	1 GYRA_STRAO	P20831 staphylococ

ALIGNMENTS

RESULT 1	TAB1_HUMAN	STANDARD;	PRT;	504 AA.
ID	TAB1_HUMAN			
AC	Q15750;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	TAK1-binding protein 1.			
GN	TAB1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID:9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE=96216294; PubMed=8638164;			
RA	Shibuya H., Yamaguchi K., Shirakabe K., Tonogawa A., Gotch Y.,			
RA	Ueno N., Irie K., Nishida E., Matsumoto K.;			
RT	"TAB1: an activator of the TAK1 MAPKK in TGF-beta signal			
RL	transduction.";			
RL	Science 272:1179-1182(1996).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20057165; PubMed=10591208;			
RA	Dunham I., Hunt A.R., Collins J.E., Bruskiewicz R., Beare D.M.,			
RA	Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,			
RA	Gilbert J.G.R., Hall-Tamlyn G., Heathcott R.W., Ho S., Holmes S.,			
RA	Hall R.S.E., Hall-Tamlyn G., Heathcott R.W., Ho S., Holmes S.,			
RA	Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,			
RA	Burill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,			
RA	Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,			
RA	Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,			
RA	Dhani P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,			
RA	Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,			
RA	Gilbert J.G.R., Hall-Tamlyn G., Heathcott R.W., Ho S., Holmes S.,			
RA	Hall R.S.E., Jones M.C., Kersey J., Kimberley A.M., King A.,			
RA	Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,			
RA	Marvin I.D., Maibreght-Mohammadi M., Matthews L.H., Mccann O.T.,			
RA	McClay J., McLaren S., McMuray A.A., Milne S.A., Morrison B.J.,			
RA	Odeh C.N., Pavlic R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,			
RA	Phillips S.H., Plumb R.W., Ramsey H., Ramsey Y., Rogers L., Ross M.T.,			
RA	Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,			
RA	Stewardland C., Spraggon L., Steward C.A., Sulston J.E., Swann R.M.,			
RA	Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,			
RA	Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,			
RA	Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,			
RA	Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,			
RA	Shirahata A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuama S.,			
RA	Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,			
RA	Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,			
RA	Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,			
RA	Pan S., Qi S., Qian Y., Ray L., Ren Q., Shauli S., Sloan D., Song L.,			
RA	Phan Q., Wang Y., Wang Z., White J., Williamson D., Wu H., Yao Z.,			
RA	Zhan M., Zhang G., Chissole S., Murray J., Miller N., Mink P.,			
RA	Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,			

```
RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
RA Hinds K., Kemp K., Latreille P., Layman D., Ozerky P., Rohlfing T.,
RA Scheet P., Walker C., Wamsley A., Wohldmann P., Pepin K., Neilson J.,
RA Korf I., Bedell J.A., Hillier L., Mardis E., Waterston R., Wilson R.,
RA Emanuel B.S., Shaikh T., Kurahashi H., Saitta S., Budarf M.L.,
RA McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E., Edelmann L.,
RA Kim U.J., Shituya H., Simon M.I., Dumaneki J.P., Peyrard M., Kedra D.,
RA Seroussi E., Fransson I., Tapia I., Bruder C.E., O'Brien K.P.,
RA Wilkenson P., Bodenteich A., Hartman K., Hu X., Khan A.S., Lane L.,
RA Tilahun Y., Wright H.;
RT "The DNA sequence of human chromosome 22.";
RL Nature 402:489-495(1999).
CC -!- FUNCTION: MAY BE AN IMPORTANT SIGNALING INTERMEDIATE BETWEEN TGFB
CC RECEPTORS AND MAP3K7/TAK1.
CC -!- SUBUNIT: Interacts with MAP3K7 and with BIRC7.
CC -!- TISSUE SPECIFICITY: UBIQUITOUS.
CC -!- SIMILARITY: CONTAINS 1 PP2C-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U49928; AAC12660.1; -.
DR EMBL; Z83845; CAB55304.1; -.
DR Genew; HGNC:18157; MAP3K7IP1.
DR MIM; 602615; -.
DR InterPro; IPR001932; PP2C-like.
DR Pfam; PF00481; PP2C, 1.
DR SMART; SM00332; PP2CC; 1.
FT DOMAIN 64 368 PP2C-LIKE.
FT DOMAIN 452 457 POLY-SER.
SQ SEQUENCE 504 AA; 54644 MW; A45743288718983A CRC64;

Query Match 100.0%; Score 359; DB 1; Length 504;
Best Local Similarity 100.0%; Pred. No. 6.9e-34;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QSTPLTQSTNTHTQSSSSDGLFRSRPAHSLPPGSDGRVEPYVDFAEYFLWSVDHG 60
Db 437 QSTPLTQSTNTHTQSSSSDGLFRSRPAHSLPPGSDGRVEPYVDFAEYFLWSVDHG 496

QY 61 EQSVWTAP 68
|||||
Db 497 EQSVWTAP 504

RESULT 2
ID IM17 ARATH STANDARD; PRT; 243 AA.
AC Q9SP15; Q9ZUS5;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Mitochondrial import inner membrane translocase subunit TIM17.
OS Arabidopsis thaliana (Mouse-ear cress).
GN Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Miernyk J.A., Coop N.E.;
RT "A component of the Arabidopsis thaliana mitochondrial inner membrane
RT protein translocase, atTIM17."
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
```

```
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Rinning C.M., Koo H.L.,
RA Moffat K.S., Cronin L.A., Shen M., Pai G., Van Aken S., Unayam L.,
RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,
RA Goodnan H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
RA Niernan W.C., White O., Eisen J.A., Salzberg S.L., Fraser C.M.,
RA Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana."
RL Nature 402:761-768(1999).
CC -!- FUNCTION: COMPONENT OF THE PREPROTEIN IMPORT MACHINERY OF THE
CC MITOCHONDRIAL INNER MEMBRANE. INTEGRAL PART OF A PROTEIN-
CC CONDUCTING CHANNEL OF THE MITOCHONDRIAL INNER MEMBRANE (BY
CC SIMILARITY).
CC -!- SUBUNIT: FORMS PART OF THE TIM23 RECEPTOR COMPLEX THAT CONSISTS OF
CC AT LEAST 3 DIFFERENT PROTEINS; TIM17, TIM23 AND TIM44 (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane (BY similarity).
CC -!- SIMILARITY: BELONGS TO THE TIM17/TIM22/TIM23 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF186847; AAF03749.1; -.
DR EMBL; AC005896; AAC98060.1; -.
DR InterPro; IPR000087; Collagen.
DR InterPro; IPR003397; Tim17.
DR Pfam; PF02466; Tim17, 1.
KW Transport; Protein transport; Translocation; Mitochondrion;
KW Inner membrane; Transmembrane.
FT TRANSMEM 15 35 POTENTIAL.
FT TRANSMEM 83 83 POTENTIAL.
FT TRANSMEM 117 137 POTENTIAL.
FT TRANSMEM 116 116 A -> T (IN REF. 1).
FT CONFLICT 219 219 Q -> H (IN REF. 1).
SQ SEQUENCE 243 AA; 25571 MW; 199285297F58BD51 CRC64;

Query Match 17.7%; Score 63.5; DB 1; Length 243;
Best Local Similarity 39.0%; Pred. No. 3.1;
Matches 16; Conservative 6; Mismatches 18; Indels 1; Gaps 1;

QY 8 QSTNTHTQSSSSDGLFRSRPAHSLPPGSDGRVEPYVDV 48
Db 193 QNONTASSSSSSWFGGLF-DKKKEVQPGSEKTEVLESF 232

RESULT 3
AD08_HUMAN
ID AD08_HUMAN STANDARD; PRT; 824 AA.
AC F78325;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE ADAM 8 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE domain 8) (Cell surface antigen MS2) (CD156a antigen) (CD156).
GN ADAM8 OR MS2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=97271556; PubMed=9126482;
RA Yoshiyama K., Higuchi Y., Kataoka M., Matsuura K., Yamamoto S.;
```

RT "CD156 human ADM8) : expression primary amino acid sequence, and
 RL gene location." ;
 CC Genomics 41:56-62 (1997).
 CC -|- FUNCTION: POSSIBLE INVOLVEMENT IN EXTRAVASATION OF LEUKOCYTES.
 CC -|- SUBCELLULAR LOCATION: type I membrane protein.
 CC -|- TISSUE SPECIFICITY: EXPRESSED ON NEUTROPHILS AND MONOCYTES.
 CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B.
 CC -|- SIMILARITY: CONTAINS 1 DISINTEGRIN DOMAIN.
 CC -|- DATABASE: NAME=PROW; NOTE=CD guide CD156 entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd156.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; D26579; BAA05626.1; -.
 DR HSSP; P18619; 1FVL.
 DR MEROPS; M12.208; -.
 DR Geneweb; HGNC:215; ADM8.
 DR MIM; 602267; .
 DR InterPro; IPR001762; Disintegrin.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR002870; Pep_M12B_propep.
 DR InterPro; IPR001590; Reprolyasin.
 DR InterPro; IPR000130; Zn_MTPepidase.
 DR Pfam; PF00200; disintegrin_1.
 DR Pfam; PF01421; Reprolyasin; 1.
 DR Pfam; PF01562; Pep_M12B_propep; 1.
 DR Pfam; PD000664; Disintegrin; 1.
 DR SMART; SMO0050; DISIN; 1.
 DR SMART; SMO0181; EGF; 1.
 DR PROSITE; PS50215; ADAM_MERO; 1.
 DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
 DR PROSITE; PS50214; DISINTEGRIN_2; 1.
 DR PROSITE; PS01186; EGF_2; UNKNOWN; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 DR K1 Hydroxylase; Metalloproteinase; zinc; Signal; Glycoprotein;
 KW Transmembrane; Antigen.
 FT SIGNAL 1 16
 FT CHAIN 17 824
 FT DOMAIN 17 655
 FT TRANSMEM 656 676
 FT DOMAIN 677 824
 FT DOMAIN 200 400
 FT DOMAIN 408 494
 FT METAL 334 334
 FT METAL 335 335
 FT ACT_SITE 335 335
 FT METAL 338 338
 FT METAL 344 344
 FT DISULFID 310 395
 FT CARBOHYD 67 91
 FT CARBOHYD 91 91
 FT CARBOHYD 436 436
 FT CARBOHYD 612 612
 FT CARBOHYD 824 AA; 88673 MW; 5DF8E05F30DF479E CRC64;
 SQ SEQUENCE
 Query Match 17.7%; Score 63.5; DB 1; Length 824;
 Best Local Similarity 38.6%; Pred. No. 13;
 Matches 27; Conservative 2; Mismatches 30; Indels 11; Gaps 3;
 QY 7 LOSTNTHQSSSSSD-----GGLFRSHPAHSLPDGEDGRVEPY--VDPAEFYRLWS-- 56
 Db 161 LQTACTGCGVSDSLGLGPRTAFAVFRPPGSLPSRRETRYELVYVVWDNAEFQMLGSA 220
 Db 221 AVRRARLEVY 230

```

RESULT 4
NU57_YEAST
ID NU57_YEAST STANDARD; PRT; 541 AA.
AC P48837;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Nucleoporin NU57 (Nuclear pore protein NU57).
GN NU57 OR YGR119C OR G6320.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Grandi P., Schlach N., Tekotte H., Hurt E.C.;
RX "Functional interaction of Nup96 with a core nucleoporin complex
RT consisting of Nup10, Nup49p and a novel protein Nup57p.";
RL EMOO J. 14:76-87(1995).
RN [2]
RP SEQUENCE OF 1-354 FROM N.A.
RC STRAIN=S288c / FY1679;
RX MEDLINE=97197982; PubMed=3046098;
RA Hansen M., Albers M., Backes U., Coblentz A., Leuther H., Neu R.,
RA Schreier A., Schaefer B., Zimmermann M., Wolf K.;
RT "The sequence of a 23.4 kb segment on the right arm of chromosome VII
RT from Saccharomyces cerevisiae reveals CLB6, SP6, RP28A and NU57
RT genes, a Ty3 element and 11 new open reading frames.";
RL Yeast 13:171-176(1997).
RN [3]
RP SEQUENCE OF 243-541 FROM N.A.
RX MEDLINE=97061913; PubMed=8905311,
RA Hansen M., Albers M., Backes U., Coblentz A., Leuther H., Neu R.,
RA Schreier A., Schaefer B., Zimmermann M., Wolf K.;
RT "The sequence of a 23.4 kb segment on the right arm of chromosome VII
RT from Saccharomyces cerevisiae reveals CLB6, SP6, RP28A and NU57
RT genes, a Ty3 element and 11 new open reading frames.";
RL Yeast 12:173-177(1996).
CC -1- FUNCTION: PART OF THE NUCLEOPORIN COMPLEX; REQUIRED FOR PROTEIN
CC TRANSPORT IN THE NUCLEUS.
CC -1- SUBUNIT: INTERACTS WITH NSP1, NUP49 AND NUP96.
CC -1- SUBCELLULAR LOCATION: Nuclear pore complex.
CC -1- DOMAIN: CONTAINS G-L-F-G REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/isb-sib.ch
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X81155; CAA57053.1; -
CC EMBL: X83039; CAA58153.1; -
CC EMBL: Z72904; CAA97129.1; -
CC EMBL: Z72905; CAA97131.1; -
CC SGD: S0003351; NUP57.
KW Nuclear protein; Transport; Coiled coil; Repeat.
FT DOMAIN 76 223 9 X 4 AA REPEATS OF G-L-F-G.
FT DOMAIN 26 31 POLY-ASN.
FT DOMAIN 127 130 POLY-THR.
FT DOMAIN 217 220 POLY-GLY.
FT DOMAIN 258 266 POLY-GLN.
FT DOMAIN 277 280 POLY-GLN.
FT DOMAIN 398 425 COILED COIL (POTENTIAL).
SQ SEQUENCE 541 AA; 57498 MW; B292AD7FB1D7E83C CRC64;

Query Match 17.3%; Score 62; DB 1; Length 541;
Best Local Similarity 50.0%; Pred. No. 12;
Matches 12; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

```

```

Db 60 QATNTFGSNQSSSTGGGLFGNKPA 83
ALSS3 CANAL STANDARD; PRT; 1119 AA.
AC 074623;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Agglutinin-like protein 3 precursor.
GN ALS3.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1161;
RX MEDLINE=98309840; PubMed=9644209;
RA Hoyer L.L., Payne T.L., Bell M., Myers A.M., Scherer S.;
RT "Candida albicans ALS3 and insights into the nature of the ALS gene family."
RL Curr. Genet. 33:451-459(1998).
RC -!- FUNCTION: MAY PLAY A ROLE IN ADHESION AND PATHOGENESIS.
CC -!- PTM: N-GLYCOSYLATED AND O-GLYCOSYLATED (POTENTIAL).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U87956; AAC39486.1; -.
KW Cell adhesion; Glycoprotein; Repeat; Signal.
FT SIGNAL 1 17
FT CHAIN 18 1119
FT DOMAIN 433 792
FT REPEAT 433 468
FT REPEAT 469 504
FT REPEAT 505 540
FT REPEAT 541 576
FT REPEAT 577 612
FT REPEAT 613 648
FT REPEAT 649 684
FT REPEAT 685 720
FT REPEAT 721 756
FT REPEAT 757 792
FT DOMAIN 399 404
FT DOMAIN 450 455
FT DOMAIN 557 563
FT DOMAIN 593 597
FT DOMAIN 630 635
FT DOMAIN 666 671
FT DOMAIN 702 707
FT DOMAIN 738 743
FT DOMAIN 774 777
FT DOMAIN 1044 1047
FT CARBOHYD 471 471
FT CARBOHYD 543 543
FT CARBOHYD 579 579
FT CARBOHYD 651 651
FT CARBOHYD 687 687
FT CARBOHYD 723 723
FT CARBOHYD 759 759
FT CARBOHYD 845 845
FT CARBOHYD 987 987
FT CARBOHYD 1050 1050
FT CARBOHYD 1061 1061
FT SEQUENCE 1119 AA; 119927 MW; 6A3FB3FC8C879A71 CRC64;

Query Match 17.0%; Score 61; DB 1; Length 1119;
Best Local Similarity 23.8%; Pred.No.38;
Matches 24; Conservative 13; Mismatches 28; Indels 36; Gaps 4;

QY 1 QSPFTLTQSTNTHTSSSSS-----SDGLFRSPAHSL----- 34
DB 720 EPPNHTVTTTWEYSQSYATTITAPPGETDTVLIREPPNHTVTTTWEYSQSYATTITII 779
QY 35 -PGEDGRV-----EPYVDFAEFYRLMSVDHGEOVVVAP 68
DB 780 APPGETDTVLIREPPNPTVTTEY---WSQSYTTATTVTAP 817

RESULT 6
PDR5_YEAST STANDARD; PRT; 1511 AA.
AC P33302;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Suppressor of toxicity of sporidesmin.
GN PDR5 OR STS1 OR YDR1 OR LEM1 OR YOR153W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AB320;
RX MEDLINE=94140838; PubMed=8307980;
RA Bissinger P.H., Kuchler K.;
RT "Molecular cloning and expression of the Saccharomyces cerevisiae
RT STS1 gene product. A yeast ABC transporter conferring mycotoxin
RT resistance."
RL J. Biol. Chem. 269:4180-4186(1994).
RN [2]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=S288c;
RX MEDLINE=94124579; PubMed=8294477;
RA Balzi E., Wang M., Leterme S., van Dyck L., Goffeau A.;
RT "PDR5, a novel yeast multidrug resistance conferring transporter
RT controlled by the transcription regulator PDR1."
RL J. Biol. Chem. 269:2206-2214(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=AH22;
RX MEDLINE=95188264; PubMed=7882421;
RA Hirata D., Yano K., Miyahara K., Miyakawa T.;
RT "Saccharomyces cerevisiae YDR1, which encodes a member of the ATP-
RT binding cassette (ABC) superfamily, is required for multidrug
RT resistance."
RL Curr. Genet. 26:285-294(1994).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / FY1678;
RX Tarasov I.A., Winsor B., Martin R.P.;
RA Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: TRANSPORTER, WHOSE PHYSIOLOGICAL FUNCTION IS NOT YET
CC ESTABLISHED. CONFERS RESISTANCE TO THE CHEMICALS CYCLOHEXIMIDE
CC AND SULFOMETHURON METHYL. EXHIBITS NUCLEOSIDE TRIPHOSPHATASE
CC ACTIVITY.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. PDR5 SUBFAMILY.
CC -!- CAUTION: REF.3 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC FRAMESHIFT IN POSITION 61.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC      or send an email to license@isb.ch).
CC
DR      EMBL; X74113; CAA52212.1; -.
DR      EMBL; L19922; AAB53769.1; -.
DR      EMBL; D26548; BAA05547.1; ALT_INIT.
DR      EMBL; U55020; AAC49639.1; -.
DR      EMBL; Z75061; CAA99359.1; -.
DR      PIR; S34702; S34702.
DR      PIR; A49730; A49730.
DR      SGD; S0005679; PDR5.
DR      InterPro; IPR003593; AAA_ATPase.
DR      InterPro; IPR003439; ABC_Transport.
DR      InterPro; IPR005285; PDR.
DR      Pfam; PF00005; ABC_tran; 2.
DR      ProDom; PD000006; ABC_transport; 2.
DR      SMART; SM00382; AAA; 1.
DR      TIGRFAMs; TIGR00956; 3a01205; 1.
DR      PROSITE; PS00211; ABC_TRANSPORTER; 1.
KM      ATP-binding; Transmembrane; Glycoprotein; Transport.
FT      DOMAIN
FT      1 517
FT      TRANSMEM 518 542
FT      TRANSMEM 559 579
FT      TRANSMEM 612 628
FT      TRANSMEM 632 650
FT      TRANSMEM 666 685
FT      TRANSMEM 775 793
FT      TRANSMEM 794 1237
FT      TRANSMEM 1238 1260
FT      TRANSMEM 1292 1313
FT      TRANSMEM 1325 1349
FT      TRANSMEM 1355 1379
FT      TRANSMEM 1389 1407
FT      TRANSMEM 1477 1499
FT      DOMAIN 1500 1511
FT      NP_BIND 905 912
FT      DOMAIN 784 787
FT      CONFLICT 171 171
FT      CONFLICT 190 190
FT      CONFLICT 214 214
FT      CONFLICT 308 308
FT      CONFLICT 340 345
FT      CONFLICT 476 476
FT      CONFLICT 648 648
FT      CONFLICT 770 770
SQ      SEQUENCE 1511 AA; 170437 MW; 4540DC0BF04744BA CRC64;

Query Match      17.0%; Score 61; DB 1; Length 1511;
Best Local Similarity 38.8%; Pred. No. 54;
Matches 19; Conservative 6; Mismatches 16; Indels 8; Gaps 3;

QY      4 TLTLSTNTHTOSSSSSDG-GLFRS-----RPAHSLP--PGEGRVPEP 44
DB      49 TLTAQSMONSTQAPKSDAIFSSGVEGVNPLFSDPAAGYDPKDDP 97

RESULT 7
OAZ_MOUSE
ID      OAZ_MOUSE          STANDARD;          PRT;          226 AA.
AC      P54369; O08610;
DT      01-OCT-1996 (Rel. 34, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Ornithine decarboxylase antizyme (ODC-Az).
GN      OAZ1 OR OAZ.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=98088944; PubMed=9428668;
RA      Nilsson J., Koskineniemi S., Persson K., Grahn B., Holm I.;
RT      "Polymamines regulate both transcription and translation of the gene

```

```

RT      encoding ornithine decarboxylase antizyme in mouse.";
RL      Eur. J. Biochem. 250:223-231(1997).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=129/SvJ;
RA      Kanare K., Uusi-Oukari M., Janne O.A.;
RL      Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
CC      -1- FUNCTION: BINDS TO, AND DESTABILIZES, ORNITHINE DECARBOXYLASE
CC      WHICH IS THEN DEGRADED. ALSO INHIBITS CELLULAR UPTAKE OF
CC      POLYAMINES BY INACTIVATING THE POLYAMINE UPTAKE TRANSPORTER.
CC      -1- MISCELLANEOUS: A RIBOSOMAL FRAMESHIFT OCCURS BETWEEN THE CODONS
CC      FOR SER-67 AND ASP-68. AN AUTOREGULATORY MECHANISM ENABLES
CC      MODULATION OF FRAMESHIFTING ACCORDING TO THE CELLULAR
CC      CONCENTRATION OF POLYAMINES.
CC      -1- SIMILARITY: BELONGS TO THE ODC ANTIZYME FAMILY.
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb.ch/announce/
CC      or send an email to license@isb.ch).
CC
CC      EMBL; U52822; AAB96329.1; -.
CC      EMBL; U52823; AAB96330.1; -.
CC      EMBL; U84291; AAC5307.1; -.
CC      WGI; WGI:109433; Oa21.
DR      InterPro; IPR02993; ODC_AZ.
DR      Pfam; PF02100; ODC_AZ; 1.
DR      ProDom; PD007483; ODC_AZ; 1.
DR      PROSITE; PS01337; ODC_AZ; 1.
FT      RIBOSOMAL frameshift.
FT      INIT MET 0
FT      CONFLICT 68 68
FT      SEQUENCE 226 AA; 25002 MW; C580D6B730318F6 CRC64;

Query Match      16.7%; Score 60; DB 1; Length 226;
Best Local Similarity 31.7%; Pred. No. 7.3;
Matches 19; Conservative 8; Mismatches 23; Indels 10; Gaps 2;

QY      1 QSRP-----LTLSNTHTQS-----SSSSDGLFRSPRPHSLPGEGRVPEVDPAE 50
DB      104 EEPSTNDKTRVLSTLTETAKQVTRPVRWVGGLYIEPLPAGLPBGSKDSFALLERPAE 163

RESULT 8
ALSL1_CANAL
ID      ALSL1_CANAL        STANDARD;          PRT;          1260 AA.
AC      P46590;
DT      01-NOV-1995 (Rel. 32, Created)
DT      01-NOV-1995 (Rel. 32, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Agglutinin-like protein 1 precursor.
GN      ALS1.
OS      Candida albicans (Yeast).
OC      Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC      Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX      NCBI_TaxID=5476;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=ATCC 11651 / B792;
RX      MEDLINE=9572392; PubMed=7752895;
RA      Hoyer L.L., Scherer S., Shatzman A.R., Livi G.P.;
RT      "Candida albicans ALS1 domains related to a Saccharomyces cerevisiae
RT      sexual agglutinin separated by a repeating motif".
RL      Mol. Microbiol. 15:39-54(1995)
CC      -1- FUNCTION: MAY PLAY A ROLE IN ADHESION AND PATHOGENESIS.
CC      -1- PTM: N-GLYCOSYLATED AND O-GLYCOSYLATED (POTENTIAL).
CC      -1- SIMILARITY: TO YEAST SAG1.
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -

```



```

CC IN THE CYTOPLASM INSTEAD OF GOING TO THE CHLOROPLAST
CC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE PLANT AND PROKARYOTIC CARBONIC
CC ANHYDRASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC DR EMBL, U08402, AAA86942.1; -
CC InterPro, IPR001765; ProX, Coanhd.
CC Pfam, PF00484; ProCA, 1.
CC DR PROSITE, PS00704; PROK_CO2_ANHYDRASE_1; 1.
CC DR PROSITE, PS00705; PROK_CO2_ANHYDRASE_2; 1.
CC Lysase; Zinc.
CC KM DOMAIN FT 1 109 CHLOROPLAST TRANSIT PEPTIDE-LIKE.
CC FT DOMAIN 41 47 POLY-SER.
CC SQ SEQUENCE 330 AA; 35547 MW; 94C7B5CGB11D768 CRC64;
CC
CC Query Match 16.4%; Score 59; DB 1; Length 330;
CC Best Local Similarity 27.0%; Pred. NO.15;
CC Matches 20; Conservative 13; Mismatches 33; Indels 8; Gaps 2;
CC
CC QY 1 GSPFTLTGISTNTHQTGSSSSSDGGLFRFRPAHSLP-----GEGDGVPEYDFAEFYR 53
CC : |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
CC 29 RGVLSAFAKFTGSSSSSSSATPPSLIRNEPVFAAPAITTPMTWEDGN-ESYEALDALK 87
CC Db 54 LMSVDHGEOSYVTA 67
CC : |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
CC 88 KWLIEKGELEPVAA 101
CC
CC RESULT 12
CC BRSL SHEEP
CC ID BRSL_SHEEP STANDARD; PRT; 399 AA.
CC AC O97967;
CC DT 30-MAY-2000 (Rel. 39, Created)
CC DT 30-MAY-2000 (Rel. 39, Last sequence update)
CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
CC DE Bombesin receptor subtype-3 (BRs-3).
CC GN BR3.
CC OS Ovis aries (Sheep).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
CC OC Bovidae; Caprinae; Ovis.
CC NCBI_TaxID=9940;
CC RN NCBI_TaxID=9940;
CC RP SEQUENCE FROM N.A.
CC RC TISSUE=Pituitary;
CC RA MEDLINE=99367324; PubMed=10425452;
CC RX Whitley J C, Moore C, Girard A S, Shulkes A;
CC RT "Molecular cloning, genomic organization and selective expression of
CC RT bombesin receptor subtype 3 in the sheep hypothalamus and
CC RT pituitary."
CC RU J. Mol. Endocrinol. 23:107-116(1999).
CC RL
CC -1- FUNCTION: ROLE IN SPERM CELL DIVISION, MATURATION, OR FUNCTION.
CC THIS RECEPTOR MEDIATES ITS ACTION BY ASSOCIATION WITH G PROTEINS
CC THAT ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER
CC SYSTEM (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```



```

OY 2 SPITLLOSTNHTGSSSSDGLFRSPAHSLPGEDEGRVPEYVDFAFYRLMS----- 56
D 388 TPNNQNTNPNCHNDTSTKST-----WLPFPGCHRRSPHYSIADVCKLKGSPEDT 437
OY 57 ----VDHGEQSV 65
D 438 LDSFLDKAEENVL 450

RESULT 14
AF17 HUMAN STANDARD; PRT; 1093 AA.
ID AF17 HUMAN
AC P55198;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE AF-17 protein.
GN MULT6 OR AF17.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=94336695; PubMed=8058765;
RA Prasad R., Leshkowitz D., Gu Y., Alder H., Nakamura T., Saito H.,
RA Huebner K., Berger R., Croce C.M., Canaan E.;
RA "Leucine-zipper dimerization motif encoded by the AF17 gene fused to
RT ALP-1 (ML) in acute leukemia.";
RT Proc. Natl. Acad. Sci. U.S.A. 91:8107-8111(1994).
RL -1- SUBCELLULAR LOCATION: Nuclear (Pentennial).
CC -1- DISEASE: INVOLVED IN ACUTE LEUKEMIAS BY A CHROMOSOMAL
CC TRANSLOCATION T(11;17)(Q23;Q21) THAT INVOLVES MLLT6 AND MLL/HRX.
CC THE RESULT IS A ROGUE ACTIVATOR PROTEIN.
CC -1- SIMILARITY: CONTAINS 2 PHD-TYPE ZINC FINGERS.
CC -1- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
CC MIM=600328;
CC WWW="http://www.infobiogen.fr/services/chronocancer/Genes/AF17.html".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U07932; AAA21145.1; -.
DR Genew; HGNC:7138; MLLT6.
DR MIM; 600328; -.
DR InterPro; IPR001965; Znf_PHD.
DR Pfam; PF00628; PHD; 1.
DR SMART; SM00249; PHD; 2.
DR PROSITE; PS01359; ZF_PHD_1; 1.
DR PROSITE; PS50016; ZF_PHD_2; 2.
KW Nuclear protein; Zinc-finger; Repeat; Chromosomal translocation;
KW Proto-oncogene.
FT ZN_FING 5 57 PHD-TYPE 1.
FT ZN_FING 51 68 C4-TYPE.
FT ZN_FING 117 180 PHD-TYPE 2.
FT DOMAIN 190 211 GLY/SER-RICH.
FT DOMAIN 275 282 POLY-SER.
FT DOMAIN 326 338 POLY-SER.
FT DOMAIN 729 764 LEUCINE-ZIPPER.
FT DOMAIN 822 829 POLY-SER.
FT DOMAIN 834 862 PRO-RICH.
FT DOMAIN 935 984 GLN-RICH.
FT DOMAIN 1040 1051 POLY-ALA.
FT DOMAIN 1069 1080 GLY-RICH.
FT SITE 551 551 MLL FUSION POINT (IN ACUTE MYELOID
LEUKEMIA PATIENT).
SQ SEQUENCE 1093 AA; 112021 MW; F60042A63BF579E CRC64;

```

```

OY 14 TOSSSSSDGLFRSPAHSLPGEDEGRVPEYVDFAFYRLMSVHGEQSVYTP 68
D 330 SSSSSSSSGGPF--QPAVS-----SLQSSPDSAFPKLQPEDEKYSKPTAP 375

Query Match 16.3%; Score 58.5; DB 1; Length 1093;
Best Local Similarity 36.4%; Pred. No. 71;
Matches 20; Conservative 8; Mismatches 18; Indels 9; Gaps 2;

RESULT 15
PLC1 CANAL STANDARD; PRT; 1099 AA.
ID PLC1 CANAL
AC O13433;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase 1
DE (EC 3.1.4.11) (PLC-1) (Phospholipase C-1).
GN PLC1.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_Taxid=5476;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=132A;
RA MEDLINE=96129081; PubMed=9467900;
RA Bennett D.E., McCreary C.E., Coleman D.C.;
RA "Genetic characterization of a phospholipase C gene from Candida
RT albicans: presence of homologous sequences in Candida species other
RT than Candida albicans.";
RT Microbiology 144:55-72(1998).
CC -1- FUNCTION: THE PRODUCTION OF THE SECOND MESSENGER MOLECULES
CC DIACYLGLYCEROL (DAG) AND INOSITOL 1,4,5-TRISPHOSPHATE (IP3) IS
CC MEDIATED BY ACTIVATED PHOSPHATIDYLINOSITOL-SPECIFIC PHOSPHOLIPASE
CC C ENZYMES.
CC -1- CATALYTIC ACTIVITY: 1-phosphatidyl-1D-myo-inositol 4,5-
CC bisphosphate + H(2)O = D-myo-inositol 1,4,5-trisphosphate +
CC diacylglycerol.
CC -1- SIMILARITY: DOMAINS X AND Y ARE CONSERVED IN DIFFERENT FORMS OF
CC PLC AND ARE ESSENTIAL FOR CATALYTIC ACTIVITY.
CC -1- SIMILARITY: CONTAINS 1 C2 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y13975; CAA74308.1; -.
DR HSRP; P10688; IDIX.
DR InterPro; IPR000008; C2.
DR InterPro; IPR001192; PI_PLC.
DR InterPro; IPR000909; PI_PLC_Xdom.
DR InterPro; IPR001711; PI_PLC_Y.
DR Pfam; PF00387; PI_PLC_X; 1.
DR Pfam; PF00389; PI_PLC_Y; 1.
DR PRINTS; PR00390; PHPLIPASEC.
DR ProDom; PD001202; PI_PLC_Y; 1.
DR SMART; SM00239; C2; 1.
DR SMART; SM00148; PLCX; 1.
DR SMART; SM00149; PLCY; 1.
DR PROSITE; PS50004; C2 DOMAIN 2; FALSE_NEG.
DR PROSITE; PS50007; PIPLC_X DOMAIN; 1.
DR PROSITE; PS50008; PIPLC_Y DOMAIN; 1.
KW Hydrolase; Lipid degradation; Transducer.
FT DOMAIN 566 726 DOMAIN X.
FT DOMAIN 794 912 DOMAIN Y.
FT DOMAIN 938 1066 C2 DOMAIN.
FT ACT_SITE 579 579 BY SIMILARITY.
FT ACT_SITE 642 642 BY SIMILARITY.

```

```
FT DOMAIN 90 102 POLY-SER.
FT DOMAIN 239 243 POLY-THR.
FT DOMAIN 517 520 POLY-ASP.
FT DOMAIN 619 623 POLY-THR.
FT DOMAIN 629 632 POLY-ASP.
FT DOMAIN 743 760 POLY-THR.
FT DOMAIN 918 921 POLY-SER.
SQ SEQUENCE 1099 AA; 124591 MW; D54D687D53A2829B CRC64;

Query Match 16.3%; Score 58.5; DB 1; Length 1099;
Best Local Similarity 28.1%; Pred.No.72;
Matches 16; Conservative 12; Mismatches 20; Indels 9; Gaps 3;

QY 1 QSPRTLQSTNTHQTSSSSSDGGLFRSR---PAHSIP--PGEDGR---VEPYVDF 48
Db 912 RKPTLKSSSNVDTRISLTINSKTIRNFELISGHQLPRFPDDYKDQAINPYSF 968

Search completed: December 10, 2002, 07:31:39
Job time : 45 secs
```



```

AC Q9CV62;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE 2310012M03Rik protein (Fragment).
GN 2310012M03Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TONGUE;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa K., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bulc D., Fleischer C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilmink L.,
RA Wynehaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayaishizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK009321; BAB26215.1;
DR MGD; MGI:1913763; 2310012M03Rik.
FT NON TER 1
SQ SEQUENCE 52 AA; 5712 MW; 813E29B1639920A6 CRC64;

Query Match 76.6%; Score 275; DB 11; Length 52;
Best Local Similarity 98.1%; Pred. No. 2.9e-26;
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 17 SSSSSDGGFLFRSPAHSLPPGSDGRVEPYVDFABFYRLWSVDHGEQSVVTAP 68
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1 SSSSSDGGFLFRSPAHSLPPGSDGRVEPYVDFABFYRLWSVDHGEQSVMTAP 52
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 3
O73614 PRELIMINARY; PRT; 498 AA.
AC O73614;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE TAB1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98130593; PubMed=9463380;
RA Shibuya H., Iwata H., Masuyama N., Gotoh Y., Yamaguchi K., Irie K.,
RA Matsumoto K., Nishida E., Ueno N.;
RA "Role of TAK1 and TAB1 in BMP signaling in early Xenopus
RT development.";
RL EMBO J. 17:1019-1028(1998).
DR EMBL; U92031; AAC14009.1;
DR InterPro; IPR001932; PP2C-like.
DR Pfam; PF00481; PP2C; 1.
DR SMART; SM00332; PP2CC; 1.

```

```

SQ SEQUENCE 498 AA; 54450 MW; 654AB8DE2353524 CRC64;

Query Match 72.6%; Score 260.5; DB 13; Length 498;
Best Local Similarity 73.5%; Pred. No. 2.4e-23;
Matches 50; Conservative 8; Mismatches 9; Indels 1; Gaps 1;

QY 1 QSTLTQLQSTNTHTQSSSSSDGGLFRSPAHSLPPGSDGRVEPYVDFABFYRLWSVDHG 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 430 QSPSATLQSTNTHTQSSSSSDGGLFRSPAHSLPPGSDGRVEPYVDFABFYRLWSVDHG 489
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 EQ-SVVTA 67
DB 490 DPGTLTLTA 497

RESULT 4
Q96S04 PRELIMINARY; PRT; 329 AA.
AC Q96S04;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypoetical 35.8 kDa protein.
GN GS78.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21096910; PubMed=1157797;
RA Daniels R.J., Peden J.F., Lloyd C., Horsley S.W., Clark K.,
RA Tufarelli C., Kearney L., Buckle V.J., Doggett N.A., Flint J.,
RA Higgs D.R.;
RT "Sequence, structure and pathology of the fully annotated terminal 2
RT Mb of the short arm of human chromosome 16.";
RL Hum. Mol. Genet. 10:339-352(2001).
DR EMBL; AE006465; AAK61262.1; -.
KW Hypoetical protein.
SQ SEQUENCE 329 AA; 35799 MW; 890FE4B3D1C5976D CRC64;

Query Match 20.8%; Score 74.5; DB 4; Length 329;
Best Local Similarity 28.6%; Pred. No. 0.65;
Matches 22; Conservative 8; Mismatches 28; Indels 19; Gaps 2;

QY 2 SPTLTQLQSTNTHTQSSSSSDGGLFRSPAHSLPPGSDG-----RVEPY 45
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 138 APTTRICRTRCHPQSGELSDG---RNRCPDASESNHGRPHGSSPVLYGYFIRICRVERN 194
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 46 VDPAFYRLWSVDHGEQ 62
DB 195 IPECEDFRTWTLGSGEK 211

RESULT 5
Q9VWB7 PRELIMINARY; PRT; 1162 AA.
AC Q9VWB7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CG11940 protein.
GN CG11940.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams W.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

```

```

RA Amandides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Blanton R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abtali J.F., Agbayan A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernier B.P., Bhandari D., Bolintineanu S.,
RA Borkova D., Botchan M.R., Bouck B., Brokstein P., Brotter P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu L.B., Davies A.,
RA Cherry J.M., Cusum M.S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Domes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Dudin K.J., Evangelista C.C., Ferrara C., Fertler S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kenison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Paclab J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitek R., Tecor C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wattam T.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RT Science 287:2185-2195 (2000).
RA EMBL: AF003513; AAF48029.1; -
DR FlyBase; FBgn0031079; CG11940.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000159; RA_domain.
DR Pfam; PF00169; PH.1.
DR SMART; SM00233; PH.1.
DR SMART; SM00314; RA.1.
DR PROSITE; PS50003; PH_DOMAIN.1.
DR PRINTS; PRO1217; PRICHEXTENSN.
SQ SEQUENCE 1162 AA; 124386 MW; CA4C252D1E9AD795 CRC64;

Query Match 19.5%; Score 70; DB 5; Length 1162;
Best Local Similarity 32.8%; Pred. No. 10;
Matches 22; Conservative 9; Mismatches 24; Indels 12; Gaps 2;

QY 3 PTLTLOSTNTHTTSSSSSDGGLFRSPRPHSLPPGDEGVVEVPFAERYRLMSVDHCE 61
DB 960 PVLFPORRPSPTLTSCHSSSSAGSAVQYAPGPMPLPFR-----ADVRLSLSSNGS 1008

QY 62 QSVYTAP 68
DB 1009 SSEVTSP 1015

RESULT 6
Q9DXA1 PRELIMINARY; PRT; 389 AA.
AC Q9DXA1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Attachment membrane glycoprotein.
GN G.
OS Avian pneumovirus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Metapneumovirus.
OX NCBI_TaxID=38525;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FR/85/1;
RX MEDLINE=20495153; PubMed=11038385;
RA Bayon-Abboyer M.H., Arnould C., Toguin D., Eterradossi N.;
RT "Nucleotide sequences of the F, L and G protein genes of two non-
RT A/non-B avian pneumoviruses (APV) reveal a novel APV subgroup.";
RL J. Gen. Virol. 81:2723-2733 (2000).
DR EMBL; AJ251085; CAC13041.1; -
DR InterPro; IPR002965; P_rich_extensn.
DR PRINTS; PRO1217; PRICHEXTENSN.
SQ SEQUENCE 389 AA; 41838 MW; 755CB4DEEC9C3EB8 CRC64;

Query Match 19.4%; Score 69.5; DB 12; Length 389;
Best Local Similarity 37.2%; Pred. No. 3.2;
Matches 16; Conservative 7; Mismatches 17; Indels 3; Gaps 1;

QY 1 QSPPTLTOSTNTHTTSSSSSDGGLFRSPRPHSLPPGDEGVVE 43
DB 274 RSPPTKRQETGRATPRNTATTOG---SSPPHSSPPGVADNME 313

RESULT 7
Q9DXA0 PRELIMINARY; PRT; 389 AA.
AC Q9DXA0;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Attachment membrane glycoprotein.
GN G.
OS Avian pneumovirus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Metapneumovirus.
OX NCBI_TaxID=38525;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FR/85/2;
RX MEDLINE=20495153; PubMed=11038385;
RA Bayon-Abboyer M.H., Arnould C., Toguin D., Eterradossi N.;
RT "Nucleotide sequences of the F, L and G protein genes of two non-
RT A/non-B avian pneumoviruses (APV) reveal a novel APV subgroup.";
RL J. Gen. Virol. 81:2723-2733 (2000).
DR EMBL; AJ288946; CAC13042.1; -
DR InterPro; IPR002965; P_rich_extensn.
DR PRINTS; PRO1217; PRICHEXTENSN.
SQ SEQUENCE 389 AA; 41812 MW; 7A54EFB8593E9F7F CRC64;

Query Match 19.4%; Score 69.5; DB 12; Length 389;
Best Local Similarity 37.2%; Pred. No. 3.2;
Matches 16; Conservative 7; Mismatches 17; Indels 3; Gaps 1;

QY 1 QSPPTLTOSTNTHTTSSSSSDGGLFRSPRPHSLPPGDEGVVE 43
DB 274 RSPPTKRQETGRATPRNTATTOG---SSPPHSSPPGVADNME 313

RESULT 8
Q9H451 PRELIMINARY; PRT; 703 AA.
AC Q9H451;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE DJ46801.1 (Atrophin 1 interacting protein 4 (AIP4)) (Fragment).
GN DJ46801.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Smith M.;

```

```

Db      119  SSTGSLPPTNTNTNTNSEGATSGLIPLTIFSGSGRPPLNPVTQAPLPFGWEQRVD----- 173
Qy      49  AEFYRLWSVDHQEQ 62
       :  |::|::|:
Db      174  -QHRVYVVDHVEK 186

RESULT 10
Q9BY75  PRELIMINARY; PRT; 862 AA.
ID      Q9BY75
AC      Q9BY75;
DT      01-JUN-2001 (TrEMBLrel. 17, Created)
DT      01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT      01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE      Ubiquitin protein ligase Itch.
GN      ITCH.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
[1]
RP      Miyazaki K., Okamoto Y., Sakamoto M., Nakagawara A.;
RL      "Homo sapiens mRNA for ubiquitin protein ligase Itch, complete cds." ;
RT      Submitted (MAR-2001) to the ENBL/GenBank/DBJ databases.
RN      [2]
SEQUENCE FROM N.A.
RP      MEDLINE=20349573; PubMed=10940313;
RX      Qiu L., Joazeiro C., Fang N., Wang H.Y., Elly C., Altman Y., Fang D.,
RA      Hunter T., Liu Y.C.;
RL      "Recognition and ubiquitination of Notch by Itch, a hect-type E3
RT      ubiquitin ligase." ;
RJ      J. Biol. Chem. 275:35734-35737(2000).
RN      [3]
SEQUENCE FROM N.A.
RP      MEDLINE=98122574; PubMed=9462742;
RX      Perry W.L., Husted C.M., Swing D.A., O'Sullivan T.N., Jenkins N.A.,
RA      Copeland N.G.;
RL      "The Itchy locus encodes a novel ubiquitin protein ligase that is
RT      disrupted in at18H mice." ;
RJ      Nat. Genet. 18:143-146(1998).
RN      [4]
SEQUENCE FROM N.A.
RP      MEDLINE=20501262; PubMed=11046148;
RX      Winberg R., Matskova L., Chen F., Plant P., Rotin D., Gish G.,
RA      Ingham B., Ernsberg I., Pawson T.;
RL      "Latent membrane protein 2A of Epstein-Barr virus binds WW domain E3
RT      protein-ubiquitin ligases that ubiquitinate B-cell tyrosine kinases." ;
RJ      Mol. Cell. Biol. 20:8526-8535(2000).
RN      [5]
SEQUENCE FROM N.A.
RP      MEDLINE=98313405; PubMed=9647693;
RX      Wood J.D., Yuan J., Margolis R.L., Colomer V., Duan K., Kushi J.,
RA      Katsinsky Z., Kleiderlein J.J., Sharp A.H., Ross C.A.;
RL      "Atrophin-1, the DRP1A gene product, interacts with two families of WW
RT      domain-containing proteins." ;
RJ      Mol. Cell. Neurosci. 11:149-160(1998).
RN      [6]
SEQUENCE FROM N.A.
RP      MEDLINE=21218930; PubMed=11318614;
RX      Chen X., Wen S.C., Fukuda M.N., Gavva N.R., Heu D.W., Akama T.O.,
RA      Yang-Peng T.L., Shen C.K.J.;
RL      "Human ITCH is a Co-Regulator of the Hematopoietic Transcription
RT      Factor NF-E2." ;
RJ      Genomics 73:238-241(2001).
RL      ENBL; AB056663; BAB39389.1; -
DR      HESL; AF095745; RAK39399.1; -
DR      HSSP; Q13526; 1PTN.
DR      InterPro; IPR000008; C2
DR      InterPro; IPR000569; HECT_domain.
DR      InterPro; IPR002349; WW.
DR      InterPro; IPR001202; WW_Rsp5_WMP.
DR      Pfam; PF00168; C21.1.

```

```

DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00397; WW; 4.
DR PRINTS; PR00403; WWDOMAIN.
DR SMART; SM00239; C2; 1.
DR SMART; SM00119; HECT; 1.
DR SMART; SM00456; WW; 4.
DR PROSITE; PS50004; C2_DOMAIN_2; 1.
DR PROSITE; PS50237; HECT; 1.
DR PROSITE; PS01159; WW_DOMAIN_1; 4.
DR PROSITE; PS50020; WW_DOMAIN_2; 2.
KW Ligase.
SQ SEQUENCE 862 AA; 98675 MW; A3D960E7F4DBF9D3 CRC64;

Query Match 18.8%; Score 67.5; DB 4; Length 862;
Best Local Similarity 31.1%; Pred. No. 14;
Matches 23; Conservative 12; Mismatches 20; Indels 19; Gaps 3;

QY 2 SPTLTQSTNTHTQSSSSSDG-----GLFRSRPAH-----SLPGEEDGRVEPYVDF 48
DB 242 SSTGSLPPTNTNTSSEGTGSLIPLTISGSGGPRPLNPVTQAPLPGEWQRVD----- 296
QY 49 AEFYRLMSVDHGEQ 62
DB 297 -QHGRVYVDHVEK 309

RESULT 11
Q96F66 PRELIMINARY; PRT; 903 AA.
ID Q96F66;
AC Q96F66;
RT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Similar to Itchy (mouse homolog) E3 ubiquitin protein
ligase.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC011571; AAH11571.1; -.
DR InterPro; IPR000008; C2.
DR InterPro; IPR000569; HECT_domain.
DR InterPro; IPR001202; WW_Rsp5_WWP.
DR Pfam; PF00168; C2; 1.
DR Pfam; PF00632; HECT; 1.
DR Pfam; PF00397; WW; 4.
DR PROSITE; PS50004; C2_DOMAIN_2; 1.
DR PROSITE; PS50237; HECT; 1.
DR PROSITE; PS01159; WW_DOMAIN_1; UNKNOWN_4.
DR PROSITE; PS50020; WW_DOMAIN_2; 2.
SQ SEQUENCE 903 AA; 102802 MW; 6777A2043C7B67BC CRC64;

Query Match 18.8%; Score 67.5; DB 4; Length 903;
Best Local Similarity 31.1%; Pred. No. 15;
Matches 23; Conservative 12; Mismatches 20; Indels 19; Gaps 3;

QY 2 SPTLTQSTNTHTQSSSSSDG-----GLFRSRPAH-----SLPGEEDGRVEPYVDF 48
DB 283 SSTGSLPPTNTNTSSEGTGSLIPLTISGSGGPRPLNPVTQAPLPGEWQRVD----- 337
QY 49 AEFYRLMSVDHGEQ 62
DB 338 -QHGRVYVDHVEK 350

RESULT 12
Q9RKD3 PRELIMINARY; PRT; 591 AA.
ID Q9RKD3

```

```

AC Q9RKD3;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Putative integral membrane protein.
GN SC03154 OR SCE87.05.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Seeger K.J., Harris D.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Thomson N.R., Parkhill J., Barrrell B.G., Rajandream M.A.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kleeser H.M., Denapalte D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleeser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kleeser T., Latke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL132674; CAB59650.1; -.
DR InterPro; IPR003342; PMT.
DR Pfam; PF02366; PMT; 1.
SQ SEQUENCE 591 AA; 66289 MW; DE7F1707C0071897 CRC64;

Query Match 18.4%; Score 66; DB 16; Length 591;
Best Local Similarity 47.4%; Pred. No. 14;
Matches 18; Conservative 2; Mismatches 14; Indels 4; Gaps 1;

QY 27 RSRPASHLPGEEDGRVEPYVDFAEFYRL----WSVDHG 60
DB 255 RSRPLAALPVDEEDGRVPRDAHVAETRLGWRPRLAAG 292

RESULT 13
Q92508 PRELIMINARY; PRT; 634 AA.
ID Q92508
AC Q92508;
RT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Dachsund-like protein DACH2.
GN DACH2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.

```

Search completed: December 10, 2002, 07:33:28
Job time : 99 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 10, 2002, 07:29:16 ; Search time 38 seconds
(without alignments)
52.652 Million cell updates/sec

Title: US-09-830-144-4_COPY_437_504

Perfect score: 359
Sequence: 1 GSPITLQSTVHTQTQSSSS.....AEFTRLMSVDHGQSVVTAP 68

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*\n2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*\n3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*\n4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*\n5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep:*\n6: /cgn2_6/ptodata/1/1aa/backfillseq1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	359	100.0	504	2	US-08-752-891-2
2	359	100.0	504	2	US-08-752-891-6
3	359	100.0	504	2	US-09-144-178-2
4	359	100.0	504	2	US-09-144-178-6
5	359	100.0	504	4	US-09-406-854-2
6	359	100.0	504	4	US-09-406-854-6
7	359	100.0	504	4	US-09-529-279-2
8	359	100.0	513	4	US-09-529-279-43
9	359	100.0	517	4	US-09-529-279-11
10	67.5	21.4	16	4	US-09-529-279-41
11	67.5	18.8	852	2	US-09-070-060-3
12	67.5	18.8	852	3	US-09-357-746-3
13	66	18.4	16	4	US-09-539-279-40
14	63.5	17.7	240	4	US-09-152-060-100
15	63.5	17.7	335	4	US-09-152-060-64
16	63.5	17.7	391	4	US-08-706-216-6
17	60.5	16.9	854	2	US-09-070-060-4
18	60.5	16.9	854	3	US-09-357-746-4
19	60	16.7	1037	4	US-09-428-711A-21
20	58.5	16.3	1093	5	US-08-545-860D-55
21	58.5	16.3	1093	5	PCT-US94-04496-55
22	57.5	16.0	421	4	US-09-322-478-6
23	57.5	16.0	856	4	US-09-152-060-77
24	56.5	15.7	922	4	US-09-116-473-4
25	56.5	15.7	1385	2	US-08-687-399-7
26	56	15.6	181	1	US-08-278-091-15
27	56	15.6	181	1	US-08-483-859-15

28	56	15.6	181	1	US-08-472-173-15	Sequence 15, Appl
29	56	15.6	181	2	US-08-487-167-15	Sequence 15, Appl
30	56	15.6	181	2	US-08-482-816-15	Sequence 15, Appl
31	56	15.6	181	2	US-08-296-148-15	Sequence 15, Appl
32	56	15.6	181	2	US-08-801-499-15	Sequence 15, Appl
33	56	15.6	181	2	US-08-615-271-15	Sequence 15, Appl
34	56	15.6	181	3	US-09-074-660-15	Sequence 15, Appl
35	56	15.6	181	3	US-09-074-659-15	Sequence 15, Appl
36	56	15.6	181	4	US-09-106-468-15	Sequence 15, Appl
37	56	15.6	181	4	US-09-106-468A-15	Sequence 15, Appl
38	56	15.6	181	4	US-09-106-467-15	Sequence 15, Appl
39	56	15.6	297	6	5514590-2	Patent No. 5514590
40	56	15.6	305	4	US-08-904-234-1	Sequence 1, Appl
41	56	15.6	315	3	US-08-965-903B-8	Sequence 8, Appl
42	56	15.6	315	4	US-09-370-398-3	Sequence 3, Appl
43	56	15.6	427	4	US-09-086-483A-4	Sequence 4, Appl
44	56	15.6	427	4	US-09-041-886-2	Sequence 2, Appl
45	56	15.6	427	4	US-09-006-353A-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-08-752-891-2 ; Application US/08752891
Sequence 2, Patent No. 5837819
GENERAL INFORMATION:
APPLICANT: MATSUMOTO, Kunihiko
APPLICANT: NISHIDA, Eisuke
TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/752,891
CLASSIFICATION: 435
FILING DATE: 20-NOV-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 8-300856
FILING DATE: 28-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 8-126282
FILING DATE: 24-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 17981/111
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO. 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 504 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-752-891-2
Query Match 100.0%; Score 359; DB 2; Length 504;
Best Local Similarity 100.0%; Pred. No. 2.3e-37;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 QSPFTLTQSTNTHTQSSSSSDGGLFRSPAHLPPGEDGRVEPYVDFAEFYRLWSVDHG 60
Db 437 QSPFTLTQSTNTHTQSSSSSDGGLFRSPAHLPPGEDGRVEPYVDFAEFYRLWSVDHG 496

Qy 61 EQSVVTAP 68
Db 497 EQSVVTAP 504

RESULT 2
US-08-752-891-6
; Sequence 6, Application US/08752891
; Patent No. 5837819
; GENERAL INFORMATION:
; APPLICANT: MATSUMOTO, Kunihiro
; APPLICANT: NISHIDA, Eisuke
; TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08752891
; FILING DATE: 20-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-300856
; FILING DATE: 28-OCT-1996
; PRIOR APPLICATION NUMBER: JP 8-126282
; FILING DATE: 24-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 17981/111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 504 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-752-891-6
Query Match 100.0%; Score 359; DB 2; Length 504;
Best Local Similarity 100.0%; Pred. No. 2.3e-37;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QSPFTLTQSTNTHTQSSSSSDGGLFRSPAHLPPGEDGRVEPYVDFAEFYRLWSVDHG 60
Db 437 QSPFTLTQSTNTHTQSSSSSDGGLFRSPAHLPPGEDGRVEPYVDFAEFYRLWSVDHG 496

Qy 61 EQSVVTAP 68
Db 497 EQSVVTAP 504

RESULT 3
US-09-144-178-2
; Sequence 2, Application US/09144178
```

```
; Patent No. 5989862
; GENERAL INFORMATION:
; APPLICANT: MATSUMOTO, Kunihiro
; APPLICANT: NISHIDA, Eisuke
; TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09144178
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/752,891
; FILING DATE: 20-NOV-1996
; APPLICATION NUMBER: JP 8-300856
; FILING DATE: 28-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-126282
; FILING DATE: 24-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 17981/111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 504 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-144-178-2
Query Match 100.0%; Score 359; DB 2; Length 504;
Best Local Similarity 100.0%; Pred. No. 2.3e-37;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QSPFTLTQSTNTHTQSSSSSDGGLFRSPAHLPPGEDGRVEPYVDFAEFYRLWSVDHG 60
Db 437 QSPFTLTQSTNTHTQSSSSSDGGLFRSPAHLPPGEDGRVEPYVDFAEFYRLWSVDHG 496

Qy 61 EQSVVTAP 68
Db 497 EQSVVTAP 504

RESULT 4
US-09-144-178-6
; Sequence 6, Application US/09144178
; Patent No. 5989862
; GENERAL INFORMATION:
; APPLICANT: MATSUMOTO, Kunihiro
; APPLICANT: NISHIDA, Eisuke
; TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
```

COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/144.178
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/752,891
FILING DATE: 20-NOV-1996
APPLICATION NUMBER: JP 8-300856
FILING DATE: 28-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 8-126282
FILING DATE: 24-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 17981/111
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 504 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-144-178-6

Query Match 100.0%; Score 359, DB 2, Length 504;
Best Local Similarity 100.0%; Pred. No. 2.3e-37;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OSPTLTQSTNTHTOSSSSSDGGLFRSRPAHSLPPGEDGRVEPYVDAEYRLMSVDHG 60
DB 437 OSPTLTQSTNTHTOSSSSSDGGLFRSRPAHSLPPGEDGRVEPYVDAEYRLMSVDHG 496

QY 61 EOSVVTAP 68
DB 497 EOSVVTAP 504

RESULT 5
US-09-406-854-2
Sequence 2, Application US/09406854
Patent No. 6140042
GENERAL INFORMATION:
APPLICANT: MATSUMOTO, Kunihito
APPLICANT: NISHIDA, Eisuke
TITLE OF INVENTION: TABI PROTEIN AND DNA CODING THEREFOR
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/406,854
FILING DATE:
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/752,891
FILING DATE: 20-NOV-1996
APPLICATION NUMBER: JP 8-300856
FILING DATE: 28-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 8-126282
FILING DATE: 24-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 17981/111
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 504 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-406-854-2

Query Match 100.0%; Score 359, DB 4, Length 504;
Best Local Similarity 100.0%; Pred. No. 2.3e-37;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OSPTLTQSTNTHTOSSSSSDGGLFRSRPAHSLPPGEDGRVEPYVDAEYRLMSVDHG 60
DB 437 OSPTLTQSTNTHTOSSSSSDGGLFRSRPAHSLPPGEDGRVEPYVDAEYRLMSVDHG 496

QY 61 EOSVVTAP 68
DB 497 EOSVVTAP 504

RESULT 6
US-09-406-854-6
Sequence 6, Application US/09406854
Patent No. 6140042
GENERAL INFORMATION:
APPLICANT: MATSUMOTO, Kunihito
APPLICANT: NISHIDA, Eisuke
TITLE OF INVENTION: TABI PROTEIN AND DNA CODING THEREFOR
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/406,854
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/752,891
FILING DATE: 20-NOV-1996
APPLICATION NUMBER: JP 8-300856
FILING DATE: 28-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 8-126282
FILING DATE: 24-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768

; REFERENCE/DOCKET NUMBER: 17981/111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 504 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-406-854-6

Query Match 100.0%; Score 359; DB 4; Length 504;
Best Local Similarity 100.0%; Pred. No. 2.3e-37;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 QSPFTLTQSTNTHTQSSSSSDGGLFRSRPAHSLPPGEDGRVEPYVDFAEFYRLWSVDHG 60
Db 437 QSPFTLTQSTNTHTQSSSSSDGGLFRSRPAHSLPPGEDGRVEPYVDFAEFYRLWSVDHG 496
Qy 61 EQSVVTAP 68
Db 497 EQSVVTAP 504

RESULT 7

US-09-529-279-2
; Sequence 2, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529,279
; CURRENT FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-279-2

Query Match 100.0%; Score 359; DB 4; Length 504;
Best Local Similarity 100.0%; Pred. No. 2.3e-37;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 QSPFTLTQSTNTHTQSSSSSDGGLFRSRPAHSLPPGEDGRVEPYVDFAEFYRLWSVDHG 60
Db 437 QSPFTLTQSTNTHTQSSSSSDGGLFRSRPAHSLPPGEDGRVEPYVDFAEFYRLWSVDHG 496
Qy 61 EQSVVTAP 68
Db 497 EQSVVTAP 504

RESULT 8

US-09-529-279-43
; Sequence 43, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278

; CURRENT APPLICATION NUMBER: US/09/529,279
; CURRENT FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 43
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-279-43

Query Match 100.0%; Score 359; DB 4; Length 513;
Best Local Similarity 100.0%; Pred. No. 2.4e-37;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 QSPFTLTQSTNTHTQSSSSSDGGLFRSRPAHSLPPGEDGRVEPYVDFAEFYRLWSVDHG 60
Db 446 QSPFTLTQSTNTHTQSSSSSDGGLFRSRPAHSLPPGEDGRVEPYVDFAEFYRLWSVDHG 505
Qy 61 EQSVVTAP 68
Db 506 EQSVVTAP 513

RESULT 9

US-09-529-279-11
; Sequence 11, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529,279
; CURRENT FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 517
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-279-11

Query Match 100.0%; Score 359; DB 4; Length 517;
Best Local Similarity 100.0%; Pred. No. 2.4e-37;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 QSPFTLTQSTNTHTQSSSSSDGGLFRSRPAHSLPPGEDGRVEPYVDFAEFYRLWSVDHG 60
Db 437 QSPFTLTQSTNTHTQSSSSSDGGLFRSRPAHSLPPGEDGRVEPYVDFAEFYRLWSVDHG 496
Qy 61 EQSVVTAP 68
Db 497 EQSVVTAP 504

RESULT 10

US-09-529-279-41
; Sequence 41, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES

```
FILE REFERENCE: 053466/0278
CURRENT APPLICATION NUMBER: US/09/529,279
CURRENT FILING DATE: 2000-04-11
PRIOR APPLICATION NUMBER: PCT/JP98/04796
PRIOR FILING DATE: 1998-10-22
PRIOR APPLICATION NUMBER: JP 9/290188
PRIOR FILING DATE: 1997-10-22
NUMBER OF SEQ ID NOS: 48
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 41
LENGTH: 16
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-529-279-41

Query Match      21.4%; Score 77; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 QSPFLTLOSTNTHQ 15
Db      2 QSPFLTLOSTNTHQ 16

RESULT 11
US-09-070-060-3
Sequence 3, Application US/09070060
Patent No. 5976849
GENERAL INFORMATION:
APPLICANT: Husted, Carolyn M.
TITLE OF INVENTION: Human E3 Ubiquitin Protein
TITLE OF INVENTION: Lysase
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESS: ZENECA Pharmaceuticals, Inc.
STREET: 1800 Concord Pike
CITY: Wilmington
STATE: DE
COUNTRY: USA
ZIP: 19850-5437
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/070,060
FILING DATE: 30-APR-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/073,839
FILING DATE: 05-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: Higgins, Patrick H
REGISTRATION NUMBER: 39,709
REFERENCE/DOCKET NUMBER: PHM.70312
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302.886.4889
TELEFAX: 302.886.8221
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 852 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-09-070-060-3
```

```
Query Match      18.8%; Score 67.5; DB 2; Length 852;
Best Local Similarity 31.1%; Pred. No. 4;
Matches 23; Conservative 12; Mismatches 20; Indels 19; Gaps 3;

QY      2 SPTLTLOSTNTHQSSSSSDG-----GLFRSRPAH-----SLPPGEDGRVEPYVDF 48
Db      232 SSTGSLPPTNTNTTSEGATGILIPLTISGSGGPRPLNPVQAPLPGEQRVD----- 286

QY      49 AEFYRLMSVDHGEQ 62
Db      287 -QHGRVYVDHVEK 299

RESULT 12
US-09-357-746-3
Sequence 3, Application US/09357746
Patent No. 6087122
GENERAL INFORMATION:
APPLICANT: ZENECA Limited
TITLE OF INVENTION: HUMAN E3 UBIQUITIN PROTEIN LIGASE
FILE REFERENCE: PHM.70312.N1
CURRENT APPLICATION NUMBER: US/09/357,746
CURRENT FILING DATE: 1999-07-21
EARLIER APPLICATION NUMBER: US No. 6087122 60/073,839
EARLIER FILING DATE: 1998-02-05
EARLIER APPLICATION NUMBER: US No. 608712209/070,060
NUMBER OF SEQ ID NOS: 15
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 3
LENGTH: 852
TYPE: PRT
ORGANISM: Homo sapiens
US-09-357-746-3

Query Match      18.8%; Score 67.5; DB 3; Length 852;
Best Local Similarity 31.1%; Pred. No. 4;
Matches 23; Conservative 12; Mismatches 20; Indels 19; Gaps 3;

QY      2 SPTLTLOSTNTHQSSSSSDG-----GLFRSRPAH-----SLPPGEDGRVEPYVDF 48
Db      232 SSTGSLPPTNTNTTSEGATGILIPLTISGSGGPRPLNPVQAPLPGEQRVD----- 286

QY      49 AEFYRLMSVDHGEQ 62
Db      287 -QHGRVYVDHVEK 299

RESULT 13
US-09-529-279-40
Sequence 40, Application US/09529279
Patent No. 6451617
GENERAL INFORMATION:
APPLICANT: ONO, KOICHIRO
APPLICANT: OHTOMO, TOSHITAKO
APPLICANT: TSUCHIYA, MASAYUKI
TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
FILE REFERENCE: 053466/0278
CURRENT APPLICATION NUMBER: US/09/529,279
CURRENT FILING DATE: 2000-04-11
PRIOR APPLICATION NUMBER: PCT/JP98/04796
PRIOR FILING DATE: 1998-10-22
PRIOR APPLICATION NUMBER: JP 9/290188
PRIOR FILING DATE: 1997-10-22
NUMBER OF SEQ ID NOS: 48
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 40
LENGTH: 16
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: peptide
```


RESULT 2
US-10-123-427-2
; Sequence 2, Application US/10123427

Patent No. US20020119525A1
GENERAL INFORMATION:
APPLICANT: MATSUMOTO, Kunihiro
NISHIDA, Eisuke
TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/123,427
FILING DATE: 17-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/406,854
FILING DATE: <Unknown>
APPLICATION NUMBER: US/08/752,891
FILING DATE: 20-NOV-1996
APPLICATION NUMBER: JP 8-300856
FILING DATE: 28-OCT-1996
APPLICATION NUMBER: JP 8-126282
FILING DATE: 24-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 17981/111
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 504 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-123-427-2

Query Match 100.0%; Score 359; DB 12; Length 504;
Best Local Similarity 100.0%; Pred. No. 1.1e-34;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QSPFTLTQSTNTHQTSSSSSDGGLFRSPAHSLPPGEDGRVEPYVDFAEFYRLWSVDHG 60
Db 437 QSPFTLTQSTNTHQTSSSSSDGGLFRSPAHSLPPGEDGRVEPYVDFAEFYRLWSVDHG 496

Qy 61 EQSVVTAP 68
Db 497 EQSVVTAP 504

RESULT 3
US-10-123-427-6
Sequence 6, Application US/10123427
Patent No. US20020119525A1
GENERAL INFORMATION:
APPLICANT: MATSUMOTO, Kunihiro
NISHIDA, Eisuke
TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500

CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/123,427
FILING DATE: 17-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/406,854
FILING DATE: <Unknown>
APPLICATION NUMBER: US/08/752,891
FILING DATE: 20-NOV-1996
APPLICATION NUMBER: JP 8-300856
FILING DATE: 28-OCT-1996
APPLICATION NUMBER: JP 8-126282
FILING DATE: 24-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 17981/111
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 504 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-123-427-6

Query Match 100.0%; Score 359; DB 12; Length 504;
Best Local Similarity 100.0%; Pred. No. 1.1e-34;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QSPFTLTQSTNTHQTSSSSSDGGLFRSPAHSLPPGEDGRVEPYVDFAEFYRLWSVDHG 60
Db 437 QSPFTLTQSTNTHQTSSSSSDGGLFRSPAHSLPPGEDGRVEPYVDFAEFYRLWSVDHG 496

Qy 61 EQSVVTAP 68
Db 497 EQSVVTAP 504

RESULT 4
US-10-158-895-43
Sequence 43, Application US/10158895
Patent No. US20020155624A1
GENERAL INFORMATION:
APPLICANT: ONO, KOICHIRO
APPLICANT: OHTOMO, TOSHIHIKO
APPLICANT: TSUCHIYA, MASAYUKI
TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
FILE REFERENCE: 053466/0278
CURRENT APPLICATION NUMBER: US/10/158,895
CURRENT FILING DATE: 2002-06-03
PRIOR APPLICATION NUMBER: US/09/529,279
PRIOR FILING DATE: 2000-04-11
PRIOR APPLICATION NUMBER: PCT/JP98/04796
PRIOR FILING DATE: 1998-10-22
PRIOR APPLICATION NUMBER: JP 9/290188
NUMBER OF SEQ ID NOS: 48
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 43


```
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-158-895-43

Query Match          100.0%; Score 359; DB 9; Length 513;
Best Local Similarity 100.0%; Pred. No. 1.1e-34;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OSPITLLOSTNTHTQSSSSSDGSLFRSRPAHSLPGEEDGRVEPYVDAEFYRLMSVDHG 60
   |||||||
Db 446 OSPITLLOSTNTHTQSSSSSDGSLFRSRPAHSLPGEEDGRVEPYVDAEFYRLMSVDHG 505
   |||||||
QY 61 EQSVVTAP 68
   |||||||
Db 506 EQSVVTAP 513
   |||||||

RESULT 5
US-10-158-895-11
; Sequence 11, Application US/10158895
; Patent No. US20020155624A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158, 895
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529, 279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 517
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-158-895-11

Query Match          100.0%; Score 359; DB 9; Length 517;
Best Local Similarity 100.0%; Pred. No. 1.1e-34;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OSPITLLOSTNTHTQSSSSSDGSLFRSRPAHSLPGEEDGRVEPYVDAEFYRLMSVDHG 60
   |||||||
Db 437 OSPITLLOSTNTHTQSSSSSDGSLFRSRPAHSLPGEEDGRVEPYVDAEFYRLMSVDHG 496
   |||||||
QY 61 EQSVVTAP 68
   |||||||
Db 497 EQSVVTAP 504
   |||||||

RESULT 6
US-09-925-300-1270
; Sequence 1270, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925, 300
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
```

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1270
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (38)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1270

Query Match          98.1%; Score 352; DB 10; Length 84;
Best Local Similarity 98.5%; Pred. No. 8.9e-35;
Matches 67; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 OSPITLLOSTNTHTQSSSSSDGSLFRSRPAHSLPGEEDGRVEPYVDAEFYRLMSVDHG 60
   |||||||
Db 17 OSPITLLOSTNTHTQSSSSSDGSLFRSRPAHSLPGEEDGRVEPYVDAEFYRLMSVDHG 76
   |||||||
QY 61 EQSVVTAP 68
   |||||||
Db 77 EQSVVTAP 84
   |||||||

RESULT 7
US-10-158-895-41
; Sequence 41, Application US/10158895
; Patent No. US20020155624A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158, 895
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529, 279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 41
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-158-895-41

Query Match          21.4%; Score 77; DB 9; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.0024;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OSPITLLOSTNTHTQ 15
   |||||||
Db 2 OSPITLLOSTNTHTQ 16
   |||||||

RESULT 8
US-09-941-831-21
; Sequence 21, Application US/09941831
; Patent No. US20020160493A1
; GENERAL INFORMATION:
; APPLICANT: Ebner et al.
; TITLE OF INVENTION: PT049P1
; FILE REFERENCE: Serine/Threonine Phosphatase Polynucleotides, Polypeptides, and Antib
; CURRENT APPLICATION NUMBER: US/09/941, 831
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: PCT/US01/06256
```

```
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/186,350
; PRIOR FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 199
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-941-831-21

Query Match      19.4%; Score 69.5; DB 9; Length 199;
Best Local Similarity 41.9%; Pred. No. 0.36;
Matches 18; Conservative 3; Mismatches 17; Indels 5; Gaps 1;

QY 1 QSPTLTLOSTNTHQTSSSSDGG-----LFRSRPAHSLPGE 38
Db 119 QSPTLTLDQTRIHSRDAFSSISGCSKFTAVRKRWDKLPVGQ 161

RESULT 9
US-10-158-895-40
; Sequence 40, Application US/10158895
; Patent No. US2002015562A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-10-158-895-40

Query Match      18.4%; Score 66; DB 9; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.047;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 VEPYVDFAEFYR 53
Db 2 VEPYVDFAEFYR 13

RESULT 10
US-10-037-667-5
; Sequence 5, Application US/10037667
; Patent No. US20020177145A1
; GENERAL INFORMATION:
; APPLICANT: Morgan, Bruce A.
; TITLE OF INVENTION: REGULATION OF NEURAL DEVELOPMENT BY
; FILE OF INVENTION: DAEDALOS
; FILE REFERENCE: 10287-044001
; CURRENT APPLICATION NUMBER: US/10/037,667
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: 60/243,110
; PRIOR FILING DATE: 2000-10-25
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0

; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/186,350
; PRIOR FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 21
; LENGTH: 199
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-037-667-5

Query Match      18.2%; Score 65.5; DB 9; Length 498;
Best Local Similarity 31.8%; Pred. No. 3.1;
Matches 28; Conservative 9; Mismatches 18; Indels 33; Gaps 5;

QY 2 SPTLTQSTNTHQTSS-----SDGGLFRSRPAHSLPGE 36
Db 372 SPTNGQDSNTDTESNHERGSGATSSRQSSAYAKEQRPDSGGLL--LFRSRMPGTAK 429

QY 37 -----CEDGRVPEPYVDFAEFYRLMSVDH 59
Db 430 SLRVLGEDG--VQVKVFKCEHCRVFLFDH 456

RESULT 11
US-09-852-797-100
; Sequence 100, Application US/09852797
; Patent No. US20020172994A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: P2003P2
; CURRENT APPLICATION NUMBER: US/09/852,797
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/265,583
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/152,060
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: PCT/US98/04858
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/040,762
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/040,710
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/050,934
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,100
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,357
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,189
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/057,765
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/048,970
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 100
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-852-797-100

Query Match      17.7%; Score 63.5; DB 9; Length 240;
Best Local Similarity 38.6%; Pred. No. 2.3;
Matches 27; Conservative 2; Mismatches 30; Indels 11; Gaps 3;

QY 7 LQSTNTHQTSSSSSD-----GGLFRSRPAHSLPGEGRVPEY--VDFAEFYRLWS-- 56
Db 47 LQTAGTCGVSDSDSLGSLGPRTAFAVFRPRGDSLPSRETRYVELVYVVDNAEFQMLGSEA 106

QY 57 -VDHGEQSVV 65
Db 107 AVNRHRLVWV 116
```

```

PQ      7  LOSTTHGSSSSSD-----GGLFRSRPHSLPPGEDGVPEY--VDFAEFYRLWS-- 56
        |||          |||          |||          |||          |||          |||
Db       47  LQTACTCGVSDSLGSLGLPRTAAVFRRPDGSJPSRRETRYVELVVWDMNMFQWLGSRA 106

OY      57  -VDHGEOSV 65
        |||          |||
DB       107 AVRHRLEV 116

RESULT 14
US-09-852-797-64
; Sequence 64, Application US/09852797
; Patent No. US20020172994A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.,
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: PZ003P2
; CURRENT APPLICATION NUMBER: US/09/852,797
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/265,583
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/152,060
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: PCT/US98/04858
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/040,762
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/040,710
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/050,934
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,100
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,357
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,189
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/057,765
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/048,970
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 100
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-852-659A-100

Query Match      17.7%; Score 63.5; DB 10; Length 240;
Best Local Similarity 38.6%; Pred.No.2.3;
Matches 27; Conservative 2; Mismatches 30; Indels 11; Gaps 3.

```

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 64
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (35)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (297)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-852-797-64

Query Match          17.7%; Score 63.5; DB 9; Length 335;
Best Local Similarity 38.6%; Pred. No. 3.3;
Matches 27; Conservative 2; Mismatches 30; Indels 11; Gaps 3;

QY 7 LQSTNTHQTQSSSSSD-----GGLFRSRPAHSLPPGEDGRVEPY--VDFAEFYRLWS-- 56
Db 161 LQTAGTCGVSDSLGSLGPRTAAVFRPRPGDSLPSRETRYVELYVVVDNAEFQMLGSEA 220

QY 57 -VDHGEQSVV 65
Db 221 AVRHRVLEVV 230

Search completed: December 10, 2002, 07:36:46
Job time : 70 secs

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-853-161-64

Query Match          17.7%; Score 63.5; DB 10; Length 335;
Best Local Similarity 38.6%; Pred. No. 3.3;
Matches 27; Conservative 2; Mismatches 30; Indels 11; Gaps 3;

QY 7 LQSTNTHQTQSSSSSD-----GGLFRSRPAHSLPPGEDGRVEPY--VDFAEFYRLWS-- 56
Db 161 LQTAGTCGVSDSLGSLGPRTAAVFRPRPGDSLPSRETRYVELYVVVDNAEFQMLGSEA 220

QY 57 -VDHGEQSVV 65
Db 221 AVRHRVLEVV 230

Search completed: December 10, 2002, 07:36:46
Job time : 70 secs

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-852-797-64

Query Match          17.7%; Score 63.5; DB 9; Length 335;
Best Local Similarity 38.6%; Pred. No. 3.3;
Matches 27; Conservative 2; Mismatches 30; Indels 11; Gaps 3;

QY 7 LQSTNTHQTQSSSSSD-----GGLFRSRPAHSLPPGEDGRVEPY--VDFAEFYRLWS-- 56
Db 161 LQTAGTCGVSDSLGSLGPRTAAVFRPRPGDSLPSRETRYVELYVVVDNAEFQMLGSEA 220

QY 57 -VDHGEQSVV 65
Db 221 AVRHRVLEVV 230

RESULT 15
US-09-853-161-64
; Sequence 64, Application US/09853161
; Patent No. US20020076756A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: P2003P3
; CURRENT APPLICATION NUMBER: US/09/853,161
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/265,583
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/152,060
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: PCT/US98/04858
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/040,762
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/040,710
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/050,934
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,100
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,357
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,189
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/057,765
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/048,970
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 64
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (35)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (297)
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 10, 2002, 04:57:56 ; Search time 3101 Seconds
(without alignments)
638.178 Million cell updates/sec

Title: US-09-830-144-4_COPY_437_504
Perfect score: 359
Sequence: 1 QSFPTLTGSTNTHGTSSSSS.....AEFTRLMSVDHGEOVVTAP 68

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Xgapop 10.0 , Xgapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 segs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODL=frame+ p2n.model -DEV=xlp
-O=/cgn2.1/USPTO.spool/US09830144/runat_04122002_141842_4974/app_query.fasta_1.263
-DB=genEmbl -QMT=fastap -STRFX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNIT=bits -START=1 -END=1 -MATRIX=bitsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09830144 @CGN 1 1 3637 @runat_04122002_141842_4974 -NCPUs=6 -ICPU=3
-NO_XLPYX -NO_MMAP -LARGOQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELDP=6 -DELEXT=7

Database : GenEmbl.*
1: gb_ba.*
2: gb_hlg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_pl.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vl.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pac.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	359	100.0	1515	6 E31041	E31041 Method for
2	359	100.0	1560	6 AR058299	AR058299 Sequence
3	359	100.0	1560	6 AR058302	AR058302 Sequence
4	359	100.0	1560	6 AR088273	AR088273 Sequence
5	359	100.0	1560	6 AR088276	AR088276 Sequence
6	359	100.0	1560	6 AR116881	AR116881 Sequence
7	359	100.0	1560	6 AR116884	AR116884 Sequence
8	359	100.0	1560	6 E14752	E14752 Human mRNA
9	359	100.0	3096	6 HS049928	U49928 Homo sapien
10	359	99.2	2944	10 BC027054	Z83845 Human DNA s
11	356	99.2	141498	2 AC127924	AC127924 Rattus no
12	356	99.2	141498	2 AC127924	AC127924 Rattus no
13	356	99.2	141498	2 AC127924	AC127924 Rattus no
14	260.5	92.6	1926	5 XLU92031	U92031 Xenopus lae
15	79	22.0	24526	2 ALB45323	ALB45323 Mus muscu
16	78	21.7	153463	2 AC115155	AC115155 Rattus no
17	76	21.2	519	6 AX386178	AX386178 Sequence
18	76	21.2	113028	9 AL365220	AL365220 Human DNA
19	75.5	21.0	82166	9 AC006237	AC006237 Homo sapi
20	75.5	21.0	179966	2 AC023247	AC023247 Homo sapi
21	75.5	21.0	199734	2 AC102141	AC102141 Mus muscu
22	75.5	21.0	205040	9 AC015849	AC015849 Homo sapi
23	75.5	21.0	241541	2 AC084085	AC084085 Homo sapi
24	75	20.9	14090	2 AC014059	AC014059 Drosophi
25	75	20.9	158402	3 AC008343	AC008343 Drosophi
26	75	20.9	25652	3 AE003811	AE003811 Drosophi
27	74	20.6	124271	2 AC123505	AC123505 Rattus no
28	74	20.6	164521	2 AC119354	AC119354 Rattus no
29	74	20.6	165265	2 AC117064	AC117064 Rattus no
30	74	20.6	171111	2 AC099452	AC099452 Rattus no
31	74	20.6	223761	2 AC099452	AC099452 Rattus no
32	74	20.6	223761	2 AC127239	AC127239 Mus muscu
33	74	20.6	317515	2 AC098888	AC098888 Mus muscu
34	73.5	20.5	169264	2 AC099415	AC099415 Mus muscu
35	73	20.3	3438	2 AC112889	AC112889 Rattus no
36	73	20.3	4440	1 ABU35611	U35611 Azospirillum
37	73	20.3	4440	1 ABU35611	U35611 Azospirillum
38	73	20.3	34266	3 LMF3391	LMF3391 Leishmani
39	73	20.3	139130	9 AC091504	AC091504 Pan trogl
40	73	20.3	163849	2 AC044808	AC044808 Mus muscu
41	72.5	20.2	226048	2 AL669881	AL669881 Mus muscu
42	72.5	20.2	131888	9 HSJ105724	AL121777 Human DNA
43	72	20.2	144877	2 AC128571	AC128571 Rattus no
44	72	20.1	10029	1 AE013381	AE013381 Methanosa
45	72	20.1	90732	9 AL670886	AL670886 Homo sapi

RESULT 1

ALIGNMENTS

Unclassified.
1 (bases 1 to 1560)
AUTHORS Matsumoto,K. and Nishida,E.
TITLE TAB1 protein
JOURNAL Patent: US 5837819-A 1 17-NOV-1998;
FEATURES
source
BASE COUNT 332 a 469 c 480 g 279 t
ORIGIN

Alignment Scores:
Pred. No.: 2,04e-30 Length: 1560
Score: 359.00 Matches: 68
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-830-144-4_COPY_437_504 (1-68) x AR058299 (1-1560)

Qy 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
Db 1338 CAAGCCGACCTTAACCTCGAGTCCACCACAGCACGACGAGCAGCAGCTCCAGC 1397

Qy 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40
Db 1398 TCTGACGAGGAGCCTTTCCGCTCCCGGCCGCCCACTCGCTCCCGCTCGCGAGACGGT 1457

Qy 41 ArgValGluProTyrValAspPheLaGluPheTyrArgLeuTrpSerValAspHisGly 60
Db 1458 CGTGTTGAGCCCTATGTGGACTTGTCTGAGTTTACCGCTCTGGAGCGTGACCATGGC 1517

Qy 61 GluGlnSerValValThrAlaPro 68
Db 1518 GAGCAGAGCGTGGTGACAGCAGC 1541

RESULT 3
AR058302
LOCUS AR058302
DEFINITION Sequence 5 from patent US 5837819.
ACCESSION AR058302
VERSION AR058302.1 GI:5983879
KEYWORDS
SOURCE
ORGANISM
Unknown.
Unclassified.
1 (bases 1 to 1560)
AUTHORS Matsumoto,K. and Nishida,E.
TITLE TAB1 protein
JOURNAL Patent: US 5837819-A 5 17-NOV-1998;
FEATURES
source
BASE COUNT 333 a 468 c 480 g 279 t
ORIGIN

Alignment Scores:
Pred. No.: 2,04e-30 Length: 1560
Score: 359.00 Matches: 68
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-830-144-4_COPY_437_504 (1-68) x AR058302 (1-1560)

Qy 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
Db 1338 CAAGCCGACCTTAACCTCGAGTCCACCACAGCACGACGAGCAGCAGCTCCAGC 1397

Qy 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40

Db 1398 TCTGACGAGGCTCTTCCGCTCCCGGCCCGCCACTCGCTCCGCTGCGAGACGCT 1457
Qy 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60
Db 1458 CGTGTGAGCCCTTAAGTGGAGCTTTGAGAGTTCAGGCTGAGCGTGGACCATGGC 1517
Qy 61 GluGlnSerValValThrAlaPro 68
Db 1518 GAGCAGACGCTGTGACGACCG 1541

RESULT 4
AR088273
LOCUS AR088273 1560 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 1 from patent US 5989862.
ACCESSION AR088273.1 GI:10015036
VERSION AR088273.1 GI:10015036
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1560)
AUTHORS Matsumoto,K. and Nishida,E.
TITLE Tab1 protein and DNA coding therefor
JOURNAL Patent: US 5989862-A 1 23-NOV-1999;
FEATURES
source location/Qualifiers
1..1560
/organism="unknown"

BASE COUNT 332 a 469 c 480 g 279 t
ORIGIN

Alignment Scores:
Pred. No.: 2.04e-30 Length: 1560
Score: 359.00 Matches: 68
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-830-144-4_COPY_437_504 (1-68) x AR088273 (1-1560)

Qy 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
Db 1338 CAAAGCCCGACCTTAACCTTGAGCTTGCAGATTTCACGACGACGACGAGCTTCAGC 1397
Qy 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGly 40
Db 1398 TCTGACGAGGCTCTTCCGCTCCCGGCCCGCCACTCGCTCCGCTGCGAGACGCT 1457
Qy 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60
Db 1458 CGTGTGAGCCCTTAAGTGGAGCTTTGAGAGTTCAGGCTGAGCGTGGACCATGGC 1517
Qy 61 GluGlnSerValValThrAlaPro 68
Db 1518 GAGCAGACGCTGTGACGACCG 1541

RESULT 5
AR088276
LOCUS AR088276 1560 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 5 from patent US 5989862.
ACCESSION AR088276
VERSION AR088276.1 GI:10015039
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1560)
AUTHORS Matsumoto,K. and Nishida,E.
TITLE Tab1 protein and DNA coding therefor
JOURNAL Patent: US 5989862-A 5 23-NOV-1999;
FEATURES
source location/Qualifiers
1..1560
/organism="unknown"

BASE COUNT 333 a 468 c 480 g 279 t
ORIGIN

Alignment Scores:
Pred. No.: 2.04e-30 Length: 1560
Score: 359.00 Matches: 68
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-830-144-4_COPY_437_504 (1-68) x AR088276 (1-1560)

Qy 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
Db 1338 CAAAGCCCGACCTTAACCTTGAGCTTGCAGATTTCACGACGACGACGAGCTTCAGC 1397
Qy 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGly 40
Db 1398 TCTGACGAGGCTCTTCCGCTCCCGGCCCGCCACTCGCTCCGCTGCGAGACGCT 1457
Qy 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60
Db 1458 CGTGTGAGCCCTTAAGTGGAGCTTTGAGAGTTCAGGCTGAGCGTGGACCATGGC 1517
Qy 61 GluGlnSerValValThrAlaPro 68
Db 1518 GAGCAGACGCTGTGACGACCG 1541

RESULT 6
AR116881
LOCUS AR116881 1560 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 1 from patent US 6140042.
ACCESSION AR116881
VERSION AR116881.1 GI:14097787
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1560)
AUTHORS Matsumoto,K. and Nishida,E.
TITLE TAB1 protein and DNA coding therefor
JOURNAL Patent: US 6140042-A 1 31-OCT-2000;
FEATURES
source location/Qualifiers
1..1560
/organism="unknown"

BASE COUNT 332 a 469 c 480 g 279 t
ORIGIN

Alignment Scores:
Pred. No.: 2.04e-30 Length: 1560
Score: 359.00 Matches: 68
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-830-144-4_COPY_437_504 (1-68) x AR116881 (1-1560)

Qy 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
Db 1338 CAAAGCCCGACCTTAACCTTGAGCTTGCAGATTTCACGACGACGACGAGCTTCAGC 1397
Qy 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGly 40
Db 1398 TCTGACGAGGCTCTTCCGCTCCCGGCCCGCCACTCGCTCCGCTGCGAGACGCT 1457
Qy 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60
Db 1458 CGTGTGAGCCCTTAAGTGGAGCTTTGAGAGTTCAGGCTGAGCGTGGACCATGGC 1517
Qy 61 GluGlnSerValValThrAlaPro 68
Db 1518 GAGCAGACGCTGTGACGACCG 1541

PC	(C12N5/10,
PC	C12R1:91), (C12P21/02, C12R1:865), (C12P21/02, C12R1:91) ; CC
strandedness:	Double;
CC	topology: Linear;
PH	Key
PH	Location/Qualifiers
FT	source
FT	1..1560
FT	/organism='Homo sapiens'
FT	/tissue_type='kidney'
FT	30..1544
FT	CDS
FT	/product='the TAB1 protein'
FT	185
FT	unsure
FT	/note='replace a'.
FEATURES	
source	Location/Qualifiers
	1..1560
	/organism="Homo sapiens"
BASE COUNT	332 a 469 c 480 g 279 t
ORIGIN	
Alignment Scores:	
Pred. No.:	2.04e-30 Length: 1560
Score:	359.00 Matches: 68
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatches: 0
Query Match:	100.00% Indels: 0
DB:	6 Gaps: 0
US-09-830-144-4_COPY_437_504 (1-68) x E14752 (1-1560)	
QY	1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
Db	1338 CAAGGCCGACCTTAACCTCGATCGATCCACCAACACGACGAGGAGGTCCTCAGC 1397
QY	21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40
Db	1398 TCTGACGAGGAGGCTCTTCGCTCCGCGCCGCGCCACTCGCTCCGCTGGCAGGACGGT 1457
QY	41 ArgValGluProTrpValAspPheAlaGluPheTyArgLeuTrpSerValAspHisGly 60
Db	1458 CGTGTTCAGCCCTATGCGACTTTGCTGAGTTTTCGCGCTCTGGAGGTGGACATGGC 1517
QY	61 GluGlnSerValValThrAlaPro 68
Db	1518 GAGCAGAGCGTGGTGACAGCACCG 1541
RESULT 9	
HSU49928	
LOCUS	HSU49928 3096 bp mRNA linear PRI 06-APR-1998
DEFINITION	Homo sapiens TAK1 binding protein (TAB1) mRNA, complete cds.
ACCESSION	U49928
VERSION	U49928.1 GI:1401125
KEYWORDS	
SOURCE	Homo sapiens.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	Shibuya,H., Yamaguchi,K., Shirakabe,K., Tonegawa,A., Gotoh,Y., Ueno,N., Irie,K., Nishida,E. and Matsumoto,K.
TITLE	TAB1: an activator of the TAK1 MAPKK in TGF-beta signal transduction
JOURNAL	Science 272 (5265), 1179-1182 (1996)
MEDLINE	96216294
PUBMED	8638164
REFERENCE	2 (bases 1 to 3096)
AUTHORS	Shibuya,H.
TITLE	Direct Submission
JOURNAL	Submitted (25-FEB-1996) Hiroshi Shibuya, Faculty of Pharmaceutical Sciences, Hokkaido University, Nishi 6-chome, Kita 12, Kita-ku, Sapporo, Hokkaido 060, Japan
FEATURES	Location/Qualifiers
source	1..3096


```
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..3096
/gene="TAK1"
21..1535
/note="TAK1"
/note="activator for TAK1"
/codon_start=1
/product="TAK1 binding protein"
/protein_id="TAK12660.1"
/db_xref="GI:1401126"
/translation="MAQRSLQSEQPSWTDPLDPLGSLGVSASNSRSADKGT
ESHPEDSMLEFSENNCPFLYGVNGYDGNRTYNAQSLSELLQLLEHAEAY
RVLLOADVVERFLESIDDALEKASLOSOLPEGPQHOLPPYQKILRLTLER
EISGSAVAVVILNKLIVYANTRALICKSTYDQVOTADHTTNEDELPLF
SOLIDAGKIKOVGILTCGSETRIDGYKYGYDIDLSAKSKPIIAPRTHGQ
PLDGYTGFVLMSBGLKYLEAHAGGQANOEIAMAIDTFKQTSLDAYQAVVDKY
KRHSIDTASGGERARCPRHEDMTLVKNFGPLGEMSOPTSPAPAPAGRVPSV
PYSSAQSTSKTSVTLSLMPSQGMVNGAHASLTDEATPTLTNOSPULTLOSTNTH
QSSSSSDGGLFRSPHSLPPEGDERVEPYDFAEFYALMSVDHSGSVVTAP"
```

BASE COUNT 642 a 936 c 952 g 566 t

ORIGIN

Alignment Scores:

Preed. No.:	4.14e-30	Length:	3096
Score:	359.00	Matches:	68
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatch:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-830-144-4_COPY_437_504 (1-68) x HSU49928 (1-3096)

QY 1 GlnSerProThrLeuThrIleuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20

Db 1329 CAAGGCCACCTTAACCTGCAGTCCACCAACGACGACGACGACGACGACCTCCACG 1388

QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGly 40

Db 1389 TCTGACGAGGAGGCTCTCCGCTCCGCGCCGCCACCTCGCTCCGCGCTGCGAGGACGGT 1448

QY 41 ArgValGluProTyrValAlaPheAlaGluPheTyrArgLeuTyrPserValaPheHisGly 60

Db 1449 CGGTTGACGCCCTATGTGACCTTCTGAGTTTACCGCCCTGAGACGTGACCATGCG 1508

QY 61 GluGlnSerValValThrAlaPro 68

Db 1509 GAGCAGAGCGGTGGTGACACGACCG 1532

RESULT 10

HS407F17

LOCUS

DEFINITION

Human DNA sequence from clone RP3-407F17 on chromosome 22 contains the gene for TAK1 (TAK1 binding protein 1), ESTs, STSs, GSSs and two putative CpG islands, complete sequence.

ACCESSION

283845

VERSION

Z83845.14 GI:5441636

KEYWORDS

HTG; CpG island; TAK1.

SOURCE

Homo sapiens

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 69660)

AUTHORS

Mamukiyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

TITLE

Direct Submission

JOURNAL

Submitted (08-DEC-1999) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk

COMMENT

Requests: clone request@sanger.ac.uk

On Jul 10, 1999 this sequence version replaced gi:5419637.

This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the

sequence is ambiguous, there is an annotation using the 'unsure' feature key.

This sequence was generated from part of bacterial clone contigs of human chromosome 22, constructed by the Sanger Centre Chromosome 22 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr22>

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP3-407F17 is from the library RPCL-3 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/> VECTOR: pCYPAC2

IMPORTANT: This sequence is not the entire insert of clone RP3-407F17. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions

The true left end of clone RP3-407F17 is at 1 in this sequence. The true left end of clone RP5-1104E15 is at 69561 in this sequence. The true right end of clone RP3-33H23 is at 17123 in this sequence. The start of this sequence overlaps with sequence AL022326.

FEATURES

source	location/Qualifiers
1..69660	/organism="Homo sapiens"
/db_xref="taxon:9606"	
/chromosome="22"	
/clone="RP3-407F17"	
/clone_1lb="RPCL-3"	
30..216	/note="AluJb repeat: matches 111..296 of consensus"
31..412	/note="match: STS: Em:HS407F17S; match: STS: Em:AL021668"
346..437	/note="MER45B repeat: matches 1..74 of consensus"
438..753	/note="AluX repeat: matches 1..308 of consensus"
754..790	/note="MER45B repeat: matches 74..114 of consensus"
879..1189	/note="AluSq repeat: matches 1..313 of consensus"
1198..1507	/note="AluSg repeat: matches 1..298 of consensus"
1855..2147	/note="AluX repeat: matches 1..293 of consensus"
2367..2845	/note="match: GSS: Em:AQ877061"
2485..2657	/note="AluSg/x repeat: matches 135..308 of consensus"
3146..3341	/note="MER45B repeat: matches 662..1039 of consensus"
3504..3817	/note="AluY repeat: matches 1..311 of consensus"
3820..3836	/note="WIR repeat: matches 193..207 of consensus"
3837..4187	/note="AluSg repeat: matches 1..312 of consensus"
4188..4356	/note="WIR repeat: matches 8..193 of consensus"
4613..4797	/note="AluSg/x repeat: matches 126..306 of consensus"
4864..5019	/note="L2 repeat: matches 1599..1763 of consensus"
5026..5206	/note="L2 repeat: matches 2241..2420 of consensus"

```
repeat_region 5293. .5594
/note="AluYb8 repeat: matches 1. .308 of consensus"
repeat_region 5597. .5890
/note="AluJb repeat: matches 12. .303 of consensus"
repeat_region 5891. .6057
/note="AluSp repeat: matches 126. .300 of consensus"
repeat_region 6068. .6236
/note="L2 repeat: matches 2544. .2750 of consensus"
repeat_region 6539. .7149
/note="L2 repeat: matches 1815. .2541 of consensus"
repeat_region 7371. .7432
/note="MIR repeat: matches 158. .219 of consensus"
repeat_region 7439. .7516
/note="AluJb repeat: matches 1. .60 of consensus"
repeat_region 7517. .7691
/note="AluSc repeat: matches 132. .306 of consensus"
repeat_region 7692. .7986
/note="AluSg repeat: matches 1. .294 of consensus"
repeat_region 7987. .8213
/note="AluJb repeat: matches 53. .281 of consensus"
repeat_region 8297. .8506
/note="AluSp repeat: matches 2. .212 of consensus"
repeat_region 8691. .8994
/note="AluX repeat: matches 1. .301 of consensus"
repeat_region 9055. .9367
/note="AluSg repeat: matches 1. .310 of consensus"
repeat_region 9532. .9655
/note="AluSg/x repeat: matches 69. .192 of consensus"
repeat_region 9656. .9804
/note="AluJb repeat: matches 149. .297 of consensus"
misc_feature 9873. .10415
/note="CpG island"
/evidence=not_experimental
gene 10049. .42131
/genes="dJ407F17.1"
mRNA join(10049. .10084,25255. .25391,25749. .25902,27040. .27126,
27960. .28098,28981. .29094,29768. .29879,32076. .32220,
36952. .37174,38270. .38432,40264. .42131)
/genes="dJ407F17.1"
/product="dJ407F17.1 (TAB1 (TAB1 binding protein 1))"
/note="match: cDNAs: Em:U92031 Em:U49928; match: ESTs:
Em:HI7295 Em:AA726609 Em:AA019355 Em:AA168567 Em:AI654435
Em:AA789149 Em:AI014848 Em:AA430502 Em:AA745822
Em:AA411895 Em:AA411894 Em:AA326751 Em:AA430621 Em:H59337
Em:AA430462 Em:AA806521 Em:R40486 Em:AA019356 Em:AA157525
Em:AA028980 Em:T34667 Em:R83147 Em:A1269211 Em:AA806517
Em:AA147685 Em:AA458466 Em:Z43346 Em:A1656463 Em:W95352
Em:AA828095 Em:R60090 Em:AA280338 Em:AA679208 Em:AA442978
Em:F08391 Em:AI085576 Em:W95394 Em:RI3055 Em:H48607
Em:R60028 Em:F02762"
/evidence=not_experimental
CDS join(10052. .10084,25255. .25391,25749. .25902,27040. .27126,
27960. .28098,28981. .29094,29768. .29879,32076. .32220,
36952. .37174,38270. .38432,40264. .40471)
/genes="dJ407F17.1"
/note="match: proteins: Tr:Q15750 Tr:O73614"
/codon_start=1
/evidence=not_experimental
/product="dJ407F17.1 (TAB1 (TAB1 binding protein 1))"
/prot_id="CAB55304.1"
/db_xref="GI:5834565"
/translation="MAAQRRLSQEQPQSWTDPLPLCHLSGVGSASNRYSADGKGT
RHVLPDSDLVKPSNNCFLYGVNGYDGNRVNFVAORLSAELLIGOLNASHADVD
EQLLQAQDVVRSFLESIDDLAAEKASQLPQGVPOHLPPOYOKTLERLKTLE
ETSGGMAVAVLLANKLYVANVTGNALLCKSTVDGLOVTLNVDPHTENDELFR
SQLGLDAGIKVGVGICQESTRIGIKYKVGYYTIDLLSAKSPFIABEIHGAQ
PLDGVTFVLVLMSEGLYKALEARHPQAGQETAAMDTEFAKQSLDRAVAQVDRV
KRHSDFASGDSKTSVTLVLYPMVSGQGVNAGSASTLDEATPTLNQSPITLQSTNTH
QSSSSSDGLFRPSRAHSLPPDGRVPEYFAEYRLWSVDHGEQSVTPAP"
10423. .10665
/note="MIR repeat: matches 3. .261 of consensus"
repeat_region 10894. .11003
/note="L2 repeat: matches 2483. .2537 of consensus"
11004. .11310
/note="AluY repeat: matches 2. .308 of consensus"
11311. .11418
/note="L2 repeat: matches 2587. .2710 of consensus"
12433. .12563
/note="FLAM C repeat: matches 2. .133 of consensus"
12622. .13098
/genes="dJ407F17.1"
/note="match: GSS: Em:AQ601543"
12823. .13100
/note="AluSg repeat: matches 1. .279 of consensus"
13182. .13227
/note="MIR repeat: matches 86. .131 of consensus"
13351. .13887
/note="L2 repeat: matches 2087. .2749 of consensus"
13888. .14192
/note="AluSg repeat: matches 1. .298 of consensus"
14193. .14515
/note="L2 repeat: matches 1790. .2087 of consensus"
14551. .14831
/note="AluU repeat: matches 1. .298 of consensus"
16401. .16671
/note="AluSg1 repeat: matches 35. .305 of consensus"
16869. .17181
/note="AluY repeat: matches 1. .310 of consensus"
17473. .17722
/note="L1 repeat: matches 4070. .4338 of consensus"
17724. .18033
/note="AluY repeat: matches 1. .308 of consensus"
18034. .18168
/note="AluY repeat: matches 1. .135 of consensus"
Alignment Scores:
Pred. No.: 1,026-28 Length: 69660
Score: 359.00 Matches: 68
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-09-830-144-4_COPY_437_504 (1-68) x HS407F17 (1-69660)
Qy 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
Db 40265 CAAAGCCGACCTTAACTCTGAGTCCACCAACGACGACGACGACGACGACG 40324
Qy 21 SerAspGlyClyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40
Db 40325 TCTGACGAGGAGGCTCTTCGCTCCGCGCCGCCACCTCGCTCCGCGAGCGGT 40384
Qy 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTyrSerValAspHisGly 60
Db 40385 CGTGTGAGCCCTATGTGAGCTTCTCGAGTTTACCGCCTCTGGAGCGTGACCATG 40444
Qy 61 GluGlnSerValThrAlaPro 68
Db 40445 GAGCAGGCGTGTGTGACGACCG 40468
RESULT 11
LOCUS BC027054 2944 bp mRNA linear ROD 07-AUG-2002
DEFINITION Mus musculus, Similar to mitogen-activated protein kinase
kinase 7 interacting protein 1, clone IMAGE:5042323, mRNA, partial
cds
ACCESSION BC027054
VERSION BC027054.1 GI:20072133
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2944)
```

AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (04-APR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT Contact: MGC help desk
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
Sequencing Center: Baylor College of Medicine Human Genome
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M., Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://lml.lnl.gov>
Series: IRAX Plate: 45 Row: K Column: 11
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Genomescan gene prediction.

FEATURES
source

CDS

1..2944
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone="IMAGE:5042323"
/tissue_type="Kidney normal, 5 month old male mouse."
/clone_1fb="NCI CGAP_Kid14"
/1ab_host="DH10B"
/note="Vector: pCMV-SPORT6"
<1..1503
/codon_start=1
/product="Similar to mitogen-activated protein kinase
kinase kinase 7 interacting protein 1"
/protein_id="AAH27054.1"
/db_xref="GI:20072134"
/translation="AQRSLQSEQSPWTDLPLCHLSGVSASNSYSADKGTES
HPEDMLKFRSENNCPFLGVNGVDGNRVNPFVAQPLSMLLQGLNTEADVR
VLLQAFDVERSEFISIDDALEAKASLOSOLPEVPOHLPPOYKILRLKALREI
SGAMAVAVILNSKLYVANYGTRALLCKSTVDGQVOTQAMNDHTMEDPLPSLO
LGLDAGKIKOMGVICGSESPRIGDYKYKQYDIDILSAKSKPILAEPTHGAPLT
DVGTPFLVMSBGLYKALEAHAGPQANOELANIDTEPAKQSLDAVAQAVDYKX
HSDTPASGSEKAPCFRHEDMTLLVKNFGPIGEMSOPTPTAPGGRVYPVSVFYS
AOSTSKTSLTSLVMPISQGVNNGSHASTLDEATPTLTIQSPFTLLQSTNTHQTSSS
SSSDGLTFSPRSHSLPGEDEGRVEPYDAFEPRLMSVDHGOSVMTAP"

BASE COUNT 676 a 835 c 863 g 570 t
ORIGIN
Alignment Scores:
Pred. NO.: 8,426-30 Length: 2944
Score: 356.00 Matches: 67
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.53% Mismatches: 0
Query Match: 99.16% Indels: 0
DB: Gaps: 0

US-09-630-144-4_COPY_437_504 (1-68) x BC027054 (1-2944)

OY 1 GlnserProthleuThrlengInserThraThHsThGInserSerSerSerSer
DB 1297 CAGAGCCCATCTGACCTGACGTCCACCAACGACCAAGAGAGAGCTCCAGC 1356
OY 21 SerAPGjlyleuPheArgSerATGProAlaHsSerleuProGjlylAAspGly 40
DB 1357 TCTGACGGGGGCTCTTCCGCTCCAGACCGGCTCACTCACTTCCACCCGAGAGAGATGGC 1416

OY 41 ArgValGluProGlyValAspPheAlaGluPheThyArgLeuTrpSerValAspHisGly 60
DB 1417 CGGTGTGAGCCCTATGTGATGACTTGTCTGAGTTTACCGACTGTGAGCGGTGACACGGC 1476
OY 61 GluGlnSerValValThrAlaPro 68
DB 1477 GACACAGCGCTATGACGCGCACT 1500

RESULT 12
AC127924/c
LOCUS
DEFINITION
ACCESSION
AC127924.1 GI:21908451
VERSION
KEYWORDS
SOURCE
ORGANISM
Rattus norvegicus.
Rattus norvegicus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE

1 (bases 1 to 141498)
Muzny, D.M., Adams, C., Adio-Odojole, B., Ali-osman, F.R., Allen, C.,
Alsbrooks, S.L., Amaralunge, H.C., Are, J.R., Ayala, M., Banks, T.,
Barbata, U., Bencon, J., Bimage, K., Blankenburg, K., Bomin, D.,
Bouck, U., Bowle, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,
Buhay, C., Burch, P., Burkelt, C., Burrell, K.L., Byrd, N.C.,
Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
Cleveland, C.D., Cox, C., Coyle, M.D., Dahorne, S.R., David, R.,
Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
Doutswalte, K.U., Drepper, H., Dugan-Kocha, S., Durbin, K.B.,
Earmhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frenzt, P.,
Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
Gorell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
Harris, C., Harris, K., Hart, M., Haylak, P., Hayes, A., Hernandez, J.,
Hernandez, O., Hodgson, A., Hognes, M., Holloway, C., Hollins, B.,
Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudan, S.,
Karlsbom, E., Kelly, S., Khan, U., King, L., Kovach, J., Kovar, C.,
Kratovic, U., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louised, H.,
Lozdo, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M.,
Miner, G., Miner, Z., Mitchell, T., Mohabhat, K., Morgan, M., Morris, S.,
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokenwo, S., Ogun, M., Okwomu, G.,
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
Rivers, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savary, G.,
Scherer, S., Scott, G., Shen, H., Shooshitari, N., Sisson, I.,
Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svaltek, A., Tabot, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansley, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Usmani, K., Vasquez, U., Vera, V., Villalob, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Waerlington, C., Watlington, S.,
Williams, G., Williamson, A., Wleczek, R., Woodson, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y.F., Zhou, D., Zorrilla, S., Nelson, D.,
Weinstock, G., and Gibbs, R.

TITLE

JOURNAL
REFERENCE
2 (bases 1 to 141498)

AUTHORS

Worley, K.C.

TITLE

JOURNAL
Submitted (19-JUN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

COMMENT

----- Genome Center
Center: Baylor College of Medicine


```

US-09-830-144-4_COPY_437_504 (1-68) x AC127924 (1-141498)
Oy      1  GlnserProThrluInserThrAsnThrHisThGlnSerSerSerSer 20
Db 125460 CAGAGCCCCCTGACCTGCGAGTCTACCAACACCAACCCAGGACGAGCTCCAGC 125401
Oy      21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProProGlyGlyAspGly 40
Db 125400 TCGGACGGGGGCTCTTCCTCCGCTCCAGACCGGCTCATCTTCACCTCGAGAGAGATGCG 125341
Oy      41 ArgValGluProGlyTyrValaAspPheAlaGluPheTyrArgLeuTyrSerValaAspHisGly 60
Db 125340 AGGGTCGAGCCCTATGTCGACTTTGCTGAGAGTTTACGACTCGAGCGCTGAGCATTGAC 125281
Oy      61 GluInserValValThrAlaPro 68
Db 125280 GAGCGAGGTGATGATGACGCGACCT 125257

RESULT 13
AC127784 176665 bp DNA linear HTG 19-JUL-2002
AC127784 Rattus norvegicus clone CH230-131B6, *** SEQUENCING IN PROGRESS
LOCUS *** 59 unordered pieces.
DEFINITION
AC127784
VERSION AC127784.1 GI:21908163
KEYWORDS HTG; HTGS-PHASE1.
SOURCE Rattus norvegicus.
ORGANISM Rattus norvegicus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
          Rattus.
REFERENCE 1 (bases 1 to 176665)
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Al-Osman,F.R., Allen,C.,
          Albrooks,S.B., Amaralunge,H.C., Aye,J.R., Ayele,M., Banks,I.,
          Barbato,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,
          Bouch,J., Bowie,S., Brieva,M., Brown,M., Bryant,N.P.,
          Butay,C., Butch,P., Butte,C., Butrell,K.L., Byrd,N.C.,
          Caron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
          Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
          Cleveland,C.D., Cox,C., Coyle,M.D., Dethorne,S.R., David,R.,
          Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
          Delaney,K.R., Delgado,O., Dem,A.L., Ding,Y., Dinh,H.H.,
          Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
          Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escoto,M.,
          Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Franz,P.,
          Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
          Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
          Harrie,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
          Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
          Homsi,F., Howard,S., Huber,U., Huiyk,S., Hume,J., Jackson,L.E.,
          Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
          Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
          Kretovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
          Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W., Louisged,H.,
          Lorado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
          Maneswati,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
          Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
          Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
          Moser,M., Neal,D., Newton,J., Newton,S.N., Nguyen,A., Nguyen,N.,
          Nguyen,N., Nickerson,E., Nwokenwo,S., Oguh,M., Okunnu,G.,
          Orangunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
          Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
          Rivers,M., Rojas,A., Rojupokan,I., Rolfe,M., Ruiz,S., Savary,G.,
          Scherger,S., Scott,G., Shen,H., Shooshari,N., Sisson,I.,
          Soedersten,E., Sonake,T., Sparks,A., Stanley,H., Stone,H.,
          Sutton,A., Syatek,A., Tabor,P., Tamezisa,A., Tamezisa,K., Tang,H.,
          Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
          Usmani,K., Vasquez,L., Vera,Y., Villalón,D., Vinson,R., Wang,Q.,
          Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
          Williams,G., Williamson,A., Wleczky,R., Wooden,S., Worley,K.,
          Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
          Weinstein,G. and Gibbs,R.
          Direct Submission

```

```

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 176665)
AUTHORS Worley,K.C.
TITLE Direct Submission
JOURNAL Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
          of Molecular and Human Genetics, Baylor College of Medicine, One
          Baylor Plaza, Houston, TX 77030, USA
          ----- Genome Center
          Center: Baylor College of Medicine
          Center code: BCM
          Web site: http://www.hgsc.bcm.tmc.edu/
          Contact: hgsc-help@bcm.tmc.edu
          ----- Project Information
          Center project name: GZWK
          Center clone name: CH230-131B6
          ----- Summary Statistics
          Sequencing vector: Plasmid;
          Chemistry: Dye-terminator Big Dye; 100% of reads
          Assembly program: Phrap; version 0.990329
          Consensus quality: 120196 bases at least Q40
          Consensus quality: 127192 bases at least Q30
          Consensus quality: 131862 bases at least Q20
          -----
          * NOTE: Estimated insert size may differ from sequence length
          * (see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 59 contigs. The true order of the pieces
          * is not known and their order in this sequence record is
          * arbitrary. Gaps between the contigs are represented as
          * runs of N, but the exact sizes of the gaps are unknown.
          * This record will be updated with the finished sequence
          * as soon as it is available and the accession number will
          * be preserved.
          *
          1 1032: contig of 1032 bp in length
          * 1033 1132: gap of unknown length
          * 1133 2439: contig of 1307 bp in length
          * 2440 2539: gap of unknown length
          * 2540 3610: contig of 1071 bp in length
          * 3611 3710: gap of unknown length
          * 3711 4955: contig of 1245 bp in length
          * 4956 5055: gap of unknown length
          * 5056 7048: contig of 1993 bp in length
          * 7049 7148: gap of unknown length
          * 7149 8902: contig of 1754 bp in length
          * 8903 9002: gap of unknown length
          * 9003 10567: contig of 1565 bp in length
          * 10568 10667: gap of unknown length
          * 10668 12258: contig of 1591 bp in length
          * 12259 12358: gap of unknown length
          * 12359 14216: contig of 1858 bp in length
          * 14217 14316: gap of unknown length
          * 14317 15760: contig of 1444 bp in length
          * 15761 15860: gap of unknown length
          * 15861 16935: contig of 1075 bp in length
          * 16936 17035: gap of unknown length
          * 17036 18984: contig of 1949 bp in length
          * 18985 18985: gap of unknown length
          * 18985 20295: contig of 1211 bp in length
          * 20296 20395: gap of unknown length
          * 20396 21672: contig of 1277 bp in length
          * 21673 21772: gap of unknown length
          * 21773 23061: contig of 1289 bp in length
          * 23062 23161: gap of unknown length
          * 23162 24596: contig of 1435 bp in length
          * 24597 24597: gap of unknown length
          * 24597 26381: contig of 1685 bp in length
          * 26382 26481: gap of unknown length
          * 26482 29059: contig of 2578 bp in length
          * 29060 29159: gap of unknown length
          * 29160 30322: contig of 1163 bp in length
          * 30323 30422: gap of unknown length
          * 30423 32594: contig of 2172 bp in length
          * 32595 32694: gap of unknown length

```

TITLE

* 160802 160901: gap of unknown length

Alignment Scores:
 Pred. No.: 5,68e-28 Length: 176665
 Score: 356.00 Matches: 67
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 98.53% Mismatches: 0
 Query Match: 99.16% Indels: 0
 DB: 2 Gaps: 0

US-09-830-144_4_COPY_437_504 (1-68) x AC127784 (1-176665)

Oy 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
 Db 163208 CAGAGCCCACTCTGACCTCGATCTTACCACACACCCACAGCAGCAGCTCCAGC 163267

Oy 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProProGlyGluaspGly 40
 Db 163268 TCGAGCGGGGGCTCTTCGGCTCAGACCGGGCTCACTCACTTCCACCTGGAGAGGATGC 163327

Oy 41 ArgValGluProTrpValAspPheAlaGluPheArgLeuTrpSerValAspHisGly 60
 Db 163328 AGGTCGACCCCTATGTGACTTTTCTGAGTTTACCGACTCTGGAGCGTGGACCATGGC 163387

Oy 61 GluGlnSerValValThrAlaPro 58
 Db 163388 GAGCAGAGTGTGATGACGGCACCT 163411

RESULT 14
 XLU92031 1926 bp mRNA linear VRT 18-APR-1998
 LOCUS Xenopus laevis TAK1 binding protein TAB1 mRNA, complete cds.
 DEFINITION
 ACCESSION U92031
 VERSION U92031.1 GI:3057037
 KEYWORDS
 SOURCE Xenopus laevis.
 ORGANISM Xenopus laevis.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 Xenopodinae; Xenopus.
 1 (bases 1 to 1926)
 Shibusawa, H., Iwata, H., Masuyama, N., Gotoh, Y., Yamaguchi, K., Irie, K.,
 Matsumoto, K., Nishida, E. and Ueno, N.
 Role of TAK1 and TAB1 in BMP signaling in early Xenopus development
 EMBO J. 17 (4), 1019-1028 (1998)
 MEDLINE 98130593
 PUBMED 9463380
 REFERENCE 2 (bases 1 to 1926)
 Shibusawa, H.
 Direct Submission
 TITLE Submitted (05-MAR-1997) Faculty of Sciences, Hokkaido University,
 Nishi 6-chome, Kita 12, Kita-ku, Sapporo, Hokkaido 060, Japan
 FEATURES
 source
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 38_1534
 /note="TAK1 binding protein"
 /codon_start=1
 /product="TAB1"
 /protein_id="AAC14009.1"
 /db_xref="GI:3057038"
 /translations="MAAPRNLLHSQSWTDDLPLCNLSGVGSASNTYNSGLGKDEH
 PYEDNMTKFRGDNNILYGVFNKXVEGTRATFVGORLAELLIGQLDPDVTDAEVRKI
 GGAMVIVLVNSKLVAVANGTNRLCKSTVDGLOVTLQNAADHTTENEDEILRLSL
 GLDPTKIKGVYIGGQSTRTGDKVYKYNFNDEILLESSTAKSPRTAEPHICQCFID
 HDTGFLVMSEGILKSAHAGPGQANOEIAAMTATEFAKQSLDEVAQALVERVKRI
 QNTSKTSVRLSLVMPGSMVNGNSSTLDGTTSTLQSPSATLQNTNTHTQSSSSSS
 DGGIFRFRSLPSIQDDEGRGVYVDFTDFVRLWNAEHNDPGLTLLTAQ"

556 a 462 c 456 g 451 t 1 others
 BASE COUNT
 ORIGIN

Alignment Scores:

Pred. No.:	1,92e-19	Length:	1926
Score:	260.50	Matches:	50
Percent Similarity:	85.29%	Conservative:	8
Best Local Similarity:	73.53%	Mismatches:	9
Query Match:	72.56%	Indels:	1
DB:	5	Gaps:	1

US-09-830-144-4_COPY_437_504 (1-68) x XU92031 (1-1926)

Qy 1 GlnserProthLeuThrLeuGlnSerThraThrHisThrGlnSerSerSer 20
 Db 1325 CAATCTCAAGTGTACTCTCCAGTCCACCAACTCAGCTCAAGAGAGATTCCAGT 1384
 Qy 21 SeraspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGly 40
 Db 1385 TCTGATGAGGCGCTGTTCCGATCTCGCCCATTCCTTTTGACAGCGATGAGAGATGGC 1444
 Qy 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60
 Db 1445 CGCGTGAACCTTACGTTGACTTCACAGACTTTTACCGGCTTTGGAAATGCAGACATTAAT 1504
 Qy 61 GluGln--SerValValThrAla 67
 Db 1505 GATCCGGGAGACTGTGCTCAGTCCA 1528

RESULT 15 AL845323 245526 bp DNA linear HTG 09-AUG-2002
 LOCUS AL845323 Mus musculus chromosome 2 clone RP23-304D11, ** SEQUENCING IN
 DEFINITION PROGRES ***, 13 unordered pieces.

ACCESSION AL845323.2 GI:22205018
 VERSION AL845323.2
 KEYWORDS HTG; HTGS_PHASE1.
 SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 245526)
 AUTHORS Plumb, B.
 TITLE Direct Submission
 JOURNAL Submitted (06-AUG-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk; clonerequests@sanger.ac.uk

COMMENT

----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: humquerry@sanger.ac.uk
 ----- Project Information
 Center project name: BM304D11
 ----- Summary Statistics
 Assembly program: XGAP4; version 4.5
 Chemistry: Dye-terminator; 100% of reads
 Consensus quality: 242595 bases at least Q40
 Consensus quality: 243391 bases at least Q30
 Consensus quality: 243858 bases at least Q20
 Insert size: 244326; sum-of-contigs
 Insert size: 235253; 1.4% error; agarose-fp
 Quality coverage: 7.70x in Q20 bases; sum-of-contigs Quality
 coverage: 8.03x in Q20 bases; agarose-fp

----- NOTE: This is a 'working draft' sequence. It currently
 * consists of 13 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 13677: contig of 13677 bp in length
 * 13678 13777: gap of 100 bp

FEATURES

source

* 13778 49716: contig of 35939 bp in length
 * 49717 49816: gap of 100 bp
 * 49817 52165: contig of 2349 bp in length
 * 52166 52265: gap of 100 bp
 * 52266 57105: contig of 4840 bp in length
 * 57106 57205: gap of 100 bp
 * 57206 64319: contig of 7114 bp in length
 * 64320 64419: gap of 100 bp
 * 64420 79204: contig of 14785 bp in length
 * 79205 79304: gap of 100 bp
 * 79305 82247: contig of 2943 bp in length
 * 82248 82347: gap of 100 bp
 * 82348 103469: contig of 21122 bp in length
 * 103470 103569: gap of 100 bp
 * 103570 194645: contig of 91076 bp in length
 * 194646 194745: gap of 100 bp
 * 194746 218401: contig of 23656 bp in length
 * 218402 218501: gap of 100 bp
 * 218502 230420: contig of 11919 bp in length
 * 230421 230520: gap of 100 bp
 * 230521 232928: contig of 2408 bp in length
 * 232929 233028: gap of 100 bp
 * 233029 245526: contig of 12498 bp in length.

Location/Qualifiers

1..245526

/organism="Mus musculus"

/db_xref="taxon:10090"

/chromosome="2"

/clone="RP23-304D11"

/clone_lib="RPCT-23"

1..13677

/note="assembly_fragment:00163"

fragment_chain:1

clone_end:r7

vector_side:left"

13778..49716

/note="assembly_fragment:00503"

fragment_chain:1"

49817..52165

/note="assembly_fragment:00985"

fragment_chain:1"

52266..57105

/note="assembly_fragment:01603"

fragment_chain:1"

57206..64319

/note="assembly_fragment:02391"

fragment_chain:1"

64420..79204

/note="assembly_fragment:00758"

fragment_chain:2"

79305..82247

/note="assembly_fragment:03362"

fragment_chain:2"

82348..103469

/note="assembly_fragment:01621"

fragment_chain:2"

103570..194645

/note="assembly_fragment:02373"

fragment_chain:2"

194746..218401

/note="assembly_fragment:01245"

fragment_chain:2"

218502..230420

/note="assembly_fragment:00911"

fragment_chain:2"

230521..232928

/note="assembly_fragment:02806"

fragment_chain:2"

233029..245526

/note="assembly_fragment:02354"

fragment_chain:2

clone_end:Spe

vector_side:right"

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 10, 2002, 04:05:41 | Search time 319 Seconds
(without alignments)
480.050 Million cell updates/sec

Title: US-09-830-144-4_COPY_437_504

Sequence: 1 QSFPTLTLQSTNTHTQSSSSS.....AEFYRLMSVDHGQSVVTAP 68

Scoring table:
BLOSUM62
Xgapod 10.0, Xgapext 0.5
Ygapod 10.0, Ygapext 0.5
Fgapod 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODL=frame+ p2n model -DRV=xip
-Q=/sgn2_1/USPTO.spool/US09830144/runat_04122002_141842_4964/app_query.fasta.1.263
-DB=Geneseq.101002 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09830144 @CGN 1.1 79 @runat 04122002 141842 4964 -NCPU=6 -ICPU=3
-NO XLPY -NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOD=10 -XGAPEXT=0.5 -FGAPOD=6 -FGAPEXT=7
-YGAPOD=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	359	100.0	1515	21	AA248861	Human TAB1 coding
2	359	100.0	1560	18	AA791175	Human TAB1 (TAK1) b
3	359	100.0	1560	18	AA791178	Human TAB1 (TAK1) b
4	359	100.0	1560	20	AA56278	Human TAB1 encodin
5	359	100.0	1560	21	AAA39106	Human TAB-1 nucleo
6	359	100.0	1569	20	AA56310	Human TAB1 encodin
7	359	100.0	1569	20	AA56282	Human TAB1-FLAG en
8	359	100.0	16877	22	ABA20494	Human nervous syst
9	359	100.0	16877	22	AA156984	Human musculoskele
10	352	98.1	696	21	AA156985	Human cancer relat
11	76	21.2	519	24	ABN61139	Drosophila melanog
12	75	20.9	14091	23	ABU11586	Drosophila melanog
13	74.5	20.8	990	23	AA580813	Human encoding novel
14	71.5	19.9	1108	14	AA049396	Class II AP endonu
15	71	19.8	996	24	ABK78928	Bacillus Clausii g
16	70	19.5	3489	23	ABU28043	Drosophila melanog
17	70	19.5	19674	23	ABU28042	Drosophila melanog
18	69.5	19.4	1492	22	AAH78731	Human HT4SG64 seri
19	69.5	19.4	3820	23	ABU12378	Drosophila melanog
20	69.5	19.4	4877	23	ABU13708	Drosophila melanog
21	69.5	19.4	9507	22	AAU07097	Human reproductive
22	69	19.2	6743	23	AA565172	DNA encoding novel
23	69	19.2	6743	23	AA588306	DNA encoding novel
24	69	19.2	12409	22	AA530238	DNA encoding rena
25	69	19.2	23580	22	AA666230	Human immune/haema
26	69	19.2	23580	22	AAK83578	Human phospholiposi
27	68.5	19.1	3068	22	AA514723	Human phospholiposi
28	68.5	19.1	3068	22	AA514724	Human phospholiposi
29	68.5	19.1	12094	23	AA595936	Protonibacterium
30	68.5	19.1	16951	22	AAH48620	Human fascin DNA f
31	68.5	19.1	16951	22	AAH48622	Human fascin DNA f
32	68	18.9	273	22	ABA70748	Human foetal liver
33	68	18.9	273	22	AAK19001	Probe #15715 for g
34	68	18.9	273	22	AAK19001	Human brain expres
35	68	18.9	273	22	AAK4951	Human bone marrow
36	68	18.9	273	22	AA125076	Probe #15009 for g
37	68	18.9	273	22	AA150921	Probe #19607 used
38	68	18.9	273	24	AB519195	Human genome-deriv
39	68	18.9	8918	22	AA526708	Human genomic DNA
40	68	18.9	8919	22	AA526709	Human genomic DNA
41	67.5	18.8	582	21	AA044783	Zea mays DNA fragm
42	67.5	18.8	2217	23	AA580664	DNA encoding novel
43	67.5	18.8	2377	22	ABA82579	Atrophin-1 interfac
44	67.5	18.8	2559	20	AA029236	Human E3 ubiquitin
45	67.5	18.8	5372	20	AA029235	Human E3 ubiquitin

ALIGNMENTS

RESULT 1
AA248861
ID AA248861 standard; CDNA; 1515 bp.
XX
XX AA248861;
AC
XX
XX 24-MAR-2000 (first entry)
DE Human TAB1 coding sequence.
XX
XX Human; TAB1, XIAP, X-linked inhibitor of apoptosis protein; TGF-beta;
KW transforming growth factor-beta activated kinase 1; monocyte migration;
KW TAK1 binding protein 1; extracellular matrix protein production;
KW cell growth inhibitor; beta-amyloid protein deposition;
KW immunosuppression; Transforming growth factor-beta; ds.
XX
XX Homo sapiens.
OS

PN JPL1326328-A.
 XX 26-NOV-1999.
 XX 13-MAY-1998; 98JP-0130378.
 XX 13-MAY-1998; 98JP-0130378.
 XX (MATS/) MATSUMOTO K.
 XX WPI; 2000-078337/07.
 DR P-PSDB; AAY59450.
 XX
 PT Screening a substance which inhibits combination of the X-linked
 PT inhibitor of apoptosis protein -
 XX
 PS Claim 2; Page 25-26; 43pp; Japanese.
 XX
 CC This sequence encodes the human TAB1 protein.
 CC The invention relates to a method for screening a substance inhibiting
 CC the formation of a complex between XIAP and TAB1, in which X-linked
 CC inhibitor of apoptosis protein (XIAP), transforming growth factor-beta
 CC activated kinase 1 (TAK1) binding protein 1 (TAB1) and a substance to be
 CC tested are contacted with each other and then the presence or formation
 CC of a complex between XIAP and TAB1 is detected. The substance can be used
 CC as a drug for extracellular matrix protein production enhancement, cell
 CC growth inhibition, monocyte migration, physiologically active substance
 CC induction, immunosuppression, and beta-amyloid protein deposition. A
 CC substance inhibiting the formation of a complex between TAB1 and XIAP as
 CC well as between XIAP and TGF-beta (transforming growth factor-beta) type
 CC I and/or type II receptor is useful as a drug.
 XX
 SQ Sequence 1515 BP; 323 A; 457 C; 463 G; 272 T; 0 other;
 Alignment Scores:
 Pred. No.: 2,336-35 Length: 1515
 Score: 359.00 Matches: 68
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 21 Gaps: 0
 US-09-830-144-4_COPY_437_504 (1-68) x AA248861 (1-1515)
 Qy 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
 Db 1309 CAAGCCCGACCTTAACCTTCAGTCCACCAACAGCAGCAGCAGCAGCTCCAGC 1368
 Qy 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40
 Db 1369 TCTGACGGAGGCGCTCTTCGCTCCCGGCCGCCCACTCGCTCCGCGAGGACGGT 1428
 Qy 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTyrSerValAspHisGly 60
 Db 1429 CGTGTTGAGCCCTATGTGACTTTCTGAGTTTACCGCTCTCGAGCGTGGACCATGGC 1488
 Qy 61 GluGlnSerValValThrAlaPro 68
 Db 1489 GAGCAGAGCGGTGTGACAGCACC 1512
 RESULT 2
 AAT91175
 ID AAT91175 standard; cDNA; 1560 BP.
 XX
 AC AAT91175;
 XX
 DT 14-APR-1998 (first entry)
 XX
 DE Human TAB1 (TAK1 binding protein) cDNA.
 XX
 KW TAB1; TAK1 binding protein; transforming growth factor-beta;
 KW signal transduction; human; ds.
 XX

OS Homo sapiens.
 XX Key Location/Qualifiers
 XX CDS 30..1544
 FT /*tag= a
 FT variation 185
 FT /*tag= b
 FT /note= "another clone has adenine at position 185,
 FT with codon AGC (Ser) altered to AGA (Arg)"
 XX
 XX EP803571-A2.
 XX PD 29-OCT-1997.
 XX
 PF 24-APR-1997; 97EP-0302808.
 XX
 PR 20-NOV-1996; 96US-0752891.
 PR 24-APR-1996; 96JP-0126282.
 PR 28-OCT-1996; 96JP-0300856.
 XX (UENO/) UENO N.
 XX
 XX Matsumoto K, Nishida E;
 XX WPI; 1997-515318/48.
 DR P-PSDB; AAW26706.
 DR
 XX DNA encoding TAK1 binding protein TAB1 - member of transforming
 PT growth factor beta receptor signal production pathway, which
 PT activates TAK-1 kinase activity upon binding
 XX
 PS Claim 1; Page 17-19; 30pp; English.
 XX
 CC This cDNA clone codes for human TAB1 (see AAW26706), a novel member
 CC of the transforming growth factor-beta receptor signal transduction
 CC pathway, which activates TAK-1 kinase activity upon binding. To
 CC obtain the full-length TAB1 coding sequence, a human kidney cDNA
 CC library was screened using as a probe a partial TAB1 cDNA obtained
 CC from a yeast two-hybrid assay for proteins that interacted with
 CC TAK1. The 5' terminus was identified by 5'RACE. 2 Different
 CC clones were sequenced, with cytosine and adenine (see AAT91178) as
 CC the 185th nucleotide, respectively, and deposited as FERM BP-5599
 CC and FERM BP-5508, respectively. Also claimed are: isolated DNA
 CC encoding a protein modified by a substitution, deletion and/or
 CC addition of 1 or more amino acids of the 504-residue TAB1 sequence;
 CC (2) DNA which can hybridise with the 1560 bp nucleic acid sequence;
 CC (3) isolated DNA encoding a protein comprising amino acids 21-579
 CC or 437-504 of the 504 TAB1 sequence; (4) DNA encoding a fusion
 CC protein comprising an above protein or polypeptide; (5) expression
 CC vector comprising an above DNA; and (6) host cell, preferably a
 CC mammalian or yeast cell, transformed by the expression vector.
 CC Cells expressing TAB1 and TAK1 can be used to screen for TGF-beta
 CC signalling pathway inhibitors by contacting the cells with a test
 CC compound, and measuring the TAK1 kinase activity.
 XX
 SQ Sequence 1560 BP; 332 A; 469 C; 480 G; 279 T; 0 other;
 Alignment Scores:
 Pred. No.: 2,426-35 Length: 1560
 Score: 359.00 Matches: 68
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 18 Gaps: 0
 US-09-830-144-4_COPY_437_504 (1-68) x AAT91175 (1-1560)
 Qy 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
 Db 1338 CAAGCCCGACCTTAACCTTCAGTCCACCAACAGCAGCAGCAGCAGCTCCAGC 1397
 Qy 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40

```

Db 1398 TCTGACGAGGAGGCTCTCCGCTCCGCGCCGCGCCACCTCGCTCCCGCTGCGAGACGCT 1457
Qy 41 ArgValAGluProTyrValAspPheAlaGluPheTyrArgLeuTyrSerValAspHisGly 60
Db 1458 CGGTGTGAGCCCTATGTGAGCTTCTGTGAGTTTACCGCTCTGAGCGGTGACCATGAC 1517
Qy 61 GluGlnSerValValThrAlaPro 68
Db 1518 GAGCAGAGCGTGTGTGACAGCACCG 1541

RESULT 3
AAAT91178
ID AAAT91178 standard; cDNA, 1560 BP.
XX
XX AAT91178;
XX
XX 14-APR-1998 (first entry)
XX
XX Human TAB1 (TAK1 binding protein) cDNA.
XX
XX TAB1; TAK1 binding protein; transforming growth factor-beta;
XX signal transduction; human; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 30..1544
XX FT /*tag= a
XX FT variation 185
XX FT /*tag= b
XX FT /note= "another clone has cytosine at position 185,
XX with codon AGA (Arg) altered to AGC (Ser)"
XX
XX PN EP803571-A2.
XX
XX PD 29-OCT-1997.
XX
XX 24-APR-1997; 97EP-0302808.
XX
XX 20-NOV-1996; 96US-0752891.
XX PR 24-APR-1996; 96JP-0126282.
XX PR 28-OCT-1996; 96JP-0300856.
XX
XX (UENO/) UENO N.
XX
XX Matsumoto K, Nishida E;
XX
XX WPI; 1997-515318/48.
XX DR P-PSDB; AAW26707.
XX
XX PT DNA encoding TAK1 binding protein TAB1 - member of transforming
XX growth factor beta receptor signal production pathway, which
XX activates TAK-1 kinase activity upon binding
XX
XX PS Example 5; Page 19-21; 30pp; English.
XX
XX CC This cDNA clone codes for human TAB1 (see AAW26707), a novel member
XX of the transforming growth factor-beta receptor signal transduction
XX pathway, which activates TAK-1 kinase activity upon binding. To
XX obtain the full-length TAB1 coding sequence, a human kidney cDNA
XX library was screened using as a probe a partial TAB1 cDNA obtained
XX from a yeast two-hybrid assay for proteins that interacted with
XX TAK1. The 5' terminus was identified by 5'RACE. 2 Different
XX clones were sequenced, with cytosine (see AAT91175) and adenine as
XX the 185th nucleotide, respectively, and deposited as FERM BP-5599
XX and FERM BP-5508, respectively. Also claimed are: isolated DNA
XX encoding a protein modified by a substitution, deletion and/or
XX addition of 1 or more amino acids of the 504-residue TAB1 sequence;
XX (2) DNA which can hybridise with the 1560 bp nucleic acid sequence;
XX (3) isolated DNA encoding a protein comprising amino acids 21-579
XX or 437-504 of the 504 TAB1 sequence; (4) DNA encoding a fusion
XX protein comprising an above protein or polypeptide; (5) expression
XX vector comprising an above DNA; and (6) host cell, preferably a

```

```

CC mammalian or yeast cell, transformed by the expression vector.
CC Cells expressing TAB1 and TAK1 can be used to screen for TGF-beta
CC signalling pathway inhibitors by contacting the cells with a test
CC compound, and measuring the TAK1 kinase activity.
XX
XX SQ Sequence 1560 BP; 333 A; 468 C; 480 G; 279 T; 0 other;

Alignment Scores:
Pred. No.: 2,42e-35 Length: 1560
Score: 359.00 Matches: 68
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0

US-09-830-144-4_copy_437_504 (1-68) x AAT91178 (1-1560)
Qy 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisGlnSerSerSerSer 20
Db 1338 CAAAGCCGACCTTAACTGTGAGTCCACACAGCAGCAGCAGCATCTCCAGC 1397
Qy 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProProGlyGlnAspGly 40
Db 1398 TCTGACGAGGAGGCTCTCCGCTCCGCGCCGCGCCACCTCGCTCCCGCTGCGAGACGCT 1457
Qy 41 ArgValAGluProTyrValAspPheAlaGluPheTyrArgLeuTyrSerValAspHisGly 60
Db 1458 CGGTGTGAGCCCTATGTGAGCTTCTGTGAGTTTACCGCTCTGAGCGGTGACCATGAC 1517
Qy 61 GluGlnSerValValThrAlaPro 68
Db 1518 GAGCAGAGCGTGTGTGACAGCACCG 1541

RESULT 4
AAKS6278
ID AAKS6278 standard; DNA, 1560 BP.
XX
XX AAKS6278;
XX
XX 21-JUL-1999 (first entry)
XX
XX Human TAB1 encoding DNA.
XX
XX Human; TAB1; screening; inhibition; TGF-beta;
XX transforming growth factor beta; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 30..1544
XX FT /*tag= a
XX
XX PN WO9921010-A1.
XX
XX PD 29-APR-1999.
XX
XX PF 22-OCT-1998; 98WO-JP04796.
XX PR 22-OCT-1997; 97JP-0230188.
XX
XX (CHUS ) CHUGAI SEIYAKU KK.
XX
XX Ohtomo T, Ono K, Tsuchiya M;
XX
XX WPI; 1999-312645/26.
XX DR P-PSDB; AAY09541.
XX
XX PT Screening for TGF-beta inhibitory substances, which are useful as
XX drugs for treatment of diseases relating to its disorder
XX
XX PS Example 1; Page 143-147; 195pp; Japanese.
XX
XX A method has been developed for screening for substances which inhibit

```

CC the binding of TAK1 polypeptide to TAB1 polypeptide. The method
 CC comprises: (a) contacting the polypeptide in the presence of a sample;
 CC and (b) detecting the amount of bound polypeptide, in which the sample
 CC can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming
 CC growth factor (TGF)-beta inhibitory substances can be used in drugs for
 CC indications e.g. as TGF-beta signal transduction inhibitors or
 CC activators, or extracellular matrix protein production enhancement
 CC inhibitors, or activators, or cell proliferation prevention inhibitors or
 CC activators, or monocytic migration inhibitors or activators, or
 CC physiological activity induction inhibitors or activators, or
 CC immunosuppression inhibitors or activators, or amyloid beta protein
 CC precipitators or activators, and such substances can also be
 CC inhibitors of the TAK1 polypeptide function, particularly kinase
 CC activity. The present sequence encodes human TAB1.

XX SQ Sequence 1560 BP; 332 A; 469 C; 480 G; 279 T; 0 other;

Alignment Scores: 2,42e-35 Length: 1560
 Pred. No.: 359.00 Matches: 68
 Score: 359.00
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 20 Gaps: 0

US-09-830-144-4_COPY_437_504 (1-68) x AAX56278 (1-1560)

Qy 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
 Db 1338 CAAGCCGACCTTAACCTCGAGTCCACCCACACGACGACGAGGAGCTCCAGC 1397
 Qy 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40
 Db 1398 TCTGACGAGGAGCTCTTCGCTCCCGCCGCCACCTCGCTCCCGCTGGGAGGACGGT 1457
 Qy 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60
 Db 1458 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTGGAGGCTGGACCATGGC 1517
 Qy 61 GluGlnSerValThrAlaPro 68
 Db 1518 GAGCAGAGCGTGGTGACAGCACCG 1541

RESULT 5

AAA39106
 ID AAA39106 standard; DNA; 1560 BP.

XX AC AAA39106;

XX DT 04-SEP-2000 (first entry)

XX DE Human TAB-1 nucleotide sequence SEQ ID NO:3.

XX KW Human; TAK-1; TAB-1; mitogen activated protein kinase; MAPK;
 KW screening; signal transduction; inhibition; inflammatory cytokine;
 KW IL-1; interleukin 1; TNF; tumour necrosis factor; inflammation;
 KW antiinflammatory; suppression; ds.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers
 XX CDS 30..1544
 XX FT /*tag= a
 XX FT /product= "TAB-1"

XX PN WO200023610-A1.

XX XX 27-APR-2000.

XX XX 21-OCT-1999; 99WO-JP05817.

XX XX 21-OCT-1998; 98JP-0299962.

PA (CHUS) CHUGAI SEIYAKU KK.
 XX PI Tsuchiya M, Ohtomo T, Sugamata Y, Matsumoto K;
 XX DR WPI; 2000-339707/29.
 XX DR P-PSDB; AAV91001.

XX Method for screening inhibitors of TAK1 signal transduction for
 PT suppression of inflammatory cytokine production and use as
 PT antiinflammatory agents

XX PS Disclosure; Page 85-90; 100pp; Japanese.

XX The present invention describes a method for screening compounds for
 CC inhibition of inflammatory cytokine signal transduction by contacting
 CC the sample with TAK1 and its receptor TAB1 and selecting for inhibition
 CC of TAK1/TAB1 binding. Also described is a method for screening compounds
 CC for inhibition of inflammatory cytokine signal transduction in which the
 CC inhibition of TAK1 phosphorylation is selected for; and drug
 CC compositions for the treatment of inflammatory disorders containing as
 CC active component an inflammatory cytokine signal transduction inhibitor.
 CC TAK1 is an essential component of the signalling process which results
 CC in release of inflammatory cytokines such as interleukin-1 (IL-1),
 CC IL-10, tumour necrosis factor (TNF) and IL-6. The methods can be used
 CC for the selection of effective antiinflammatory agents. The present
 CC sequence encodes human TAB-1, which is used in the exemplification of the
 CC present invention.

XX SQ Sequence 1560 BP; 332 A; 469 C; 480 G; 279 T; 0 other;

Alignment Scores:

Pred. No.: 2,42e-35 Length: 1560
 Score: 359.00 Matches: 68
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 21 Gaps: 0

US-09-830-144-4_COPY_437_504 (1-68) x AAA39106 (1-1560)

Qy 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
 Db 1338 CAAGCCGACCTTAACCTCGAGTCCACCCACACGACGACGAGGAGCTCCAGC 1397

Qy 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40
 Db 1398 TCTGACGAGGAGCTCTTCGCTCCCGCCGCCACCTCGCTCCCGCTGGGAGGACGGT 1457

Qy 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60
 Db 1458 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCTCTGGAGGCTGGACCATGGC 1517

Qy 61 GluGlnSerValThrAlaPro 68

Db 1518 GAGCAGAGCGTGGTGACAGCACCG 1541

RESULT 6

AAA56310

ID AAA56310 standard; DNA; 1568 BP.

XX AC AAA56310;

XX DT 21-JUL-1999 (first entry)

XX DE Human TAB1 encoding DNA SEQ ID NO:42.

XX KW Human; TAB1; TAK1; screening; inhibition; TGF-beta;
 KW transforming growth factor beta; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers
 XX CDS 11..1552

```
FT XX /+tag= a
XX PN WO921010-A1.
XX PD 29-APR-1999.
XX PF 22-OCT-1998; 98WO-JP04796.
XX PR 22-OCT-1997; 97JP-0290188.
XX PA (CHUS ) CHUGAI SEIYAKU KK.
XX PI Ohtomo T, Ono K, Tsuchiya M;
XX DR WPI; 1999-312645/26.
XX DR P-PSDB; AAY09550.
XX PT Screening for TGF- beta inhibitory substances, which are useful as
XX PT drugs for treatment of diseases relating to its disorder
XX PS Example 13; Page 182-186; 195pp; Japanese.
XX CC A method has been developed for screening for substances which inhibit
XX CC the binding of TAK1 polypeptide to TAB1 polypeptide. The method
XX CC comprises: (a) contacting the polypeptide in the presence of a sample;
XX CC and (b) detecting the amount of bound polypeptide, in which the sample
XX CC can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming
XX CC growth factor (TGF)-beta inhibitory substances can be used in drugs for
XX CC indications e.g. as TGF-beta signal transduction inhibitors or
XX CC activators, or extracellular matrix protein production enhancement
XX CC inhibitors, or monocytic migration inhibitors or activators, or
XX CC physiological activity induction inhibitors or activators, or
XX CC immunosuppression inhibitors or activators, and such substances can also be
XX CC precipitation inhibitors or activators, and such substances can also be
XX CC inhibitors of the TAK1 polypeptide function, particularly kinase
XX CC activity. The present sequence encodes human TAB1.
XX SQ Sequence 1568 BP; 339 A; 472 C; 477 G; 280 T; 0 other;

Alignment Scores:
Pred. No.: 2,44e-35 Length: 1568
Score: 359.00 Matches: 68
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-830-144-4_COPY_437_504 (1-68) x AAX56310 (1-1568)
QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
Db 1346 CAAAGCCCGACCTTAACCTTCGACGTCCACCAACGACACGACGAGAGCTCCAGC 1405
QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGlyAspGly 40
Db 1406 TCTGACGAGAGCGCTTCCTCCGCTCCGCGCCCACTCGCTCCGCTGGGAGAGAGGT 1465
QY 41 ArgValGlnProTyrValAspPheAlaGlnPheTyrArgLeuThrSerValAspHisGly 60
Db 1466 CGTGTGAGCCCTATGTGAGCTTTCCTGAGCTTTACCGGCTCTGAGAGCTGACCAATGCC 1525
QY 61 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 68
Db 1526 GAGCAGAGCGGTGTGACAGACCG 1549

RESULT 7
AAX56282
ID AAX56282 standard; DNA; 1569 BP.
XX AC AAX56282;
XX XX
DT 21-JUN-1999 (first entry)
```

```
XX XX Human TAB1-FLAG encoding DNA.
XX DE
XX XX Human; TAB1; screening; inhibition; TGF-beta;
XX KM transforming growth factor beta; ss.
XX KW
XX OS Homo sapiens.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT CDS 7..1560
XX FT /+tag= a
XX PN WO921010-A1.
XX PD 29-APR-1999.
XX PF 22-OCT-1998; 98WO-JP04796.
XX PR 22-OCT-1997; 97JP-0290188.
XX PA (CHUS ) CHUGAI SEIYAKU KK.
XX PI Ohtomo T, Ono K, Tsuchiya M;
XX DR WPI; 1999-312645/26.
XX DR P-PSDB; AAY09546.
XX PT Screening for TGF- beta inhibitory substances, which are useful as
XX PT drugs for treatment of diseases relating to its disorder
XX PS Example 1; Page 159-163; 195pp; Japanese.
XX CC A method has been developed for screening for substances which inhibit
XX CC the binding of TAK1 polypeptide to TAB1 polypeptide. The method
XX CC comprises: (a) contacting the polypeptide in the presence of a sample;
XX CC and (b) detecting the amount of bound polypeptide, in which the sample
XX CC can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming
XX CC growth factor (TGF)-beta inhibitory substances can be used in drugs for
XX CC indications e.g. as TGF-beta signal transduction inhibitors or
XX CC activators, or extracellular matrix protein production enhancement
XX CC inhibitors, or monocytic migration inhibitors or activators, or
XX CC physiological activity induction inhibitors or activators, or
XX CC immunosuppression inhibitors or activators, and such substances can also be
XX CC precipitation inhibitors or activators, and such substances can also be
XX CC inhibitors of the TAK1 polypeptide function, particularly kinase
XX CC activity. The present sequence encodes TAB1-FLAG from an example of
XX CC the present invention.
XX SQ Sequence 1569 BP; 343 A; 466 C; 476 G; 284 T; 0 other;

Alignment Scores:
Pred. No.: 2,44e-35 Length: 1569
Score: 359.00 Matches: 68
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-830-144-4_COPY_437_504 (1-68) x AAX56282 (1-1569)
QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
Db 1315 CAAAGCCCGACCTTAACCTTCGACGTCCACCAACGACACGAGACACAGCTCCAGC 1374
QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGlyAspGly 40
Db 1375 TCTGACGAGAGCGCTTCCTCCGCTCCGCGCCCACTCGCTCCGCTGGGAGAGAGGT 1434
QY 41 ArgValGlnProTyrValAspPheAlaGlnPheTyrArgLeuThrSerValAspHisGly 60
Db 1435 CGTGTGAGCCCTATGTGAGCTTTCCTGAGCTTTACCGGCTCTGAGAGCTGACCAATGCC 1494
```

QY 61 GluGlnSerValValThrAlaPro 68
Db 1495 GAGCAGCGTGGTGACGACCG 1518
RESULT 8
ABA20494
ID ABA20494 standard; DNA; 16877 BP.
XX
AC ABA20494;
XX
DT 23-JAN-2002 (first entry)
XX
DE Human nervous system related polynucleotide SEQ ID NO 12825.
XX
KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX
OS Homo sapiens.
XX
PN WO200159063-A2.
XX
PD 16-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01334.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 23-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 29-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 20-OCT-2000; 2000US-0242221.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.

PR 17-NOV-2000; 2000US-0249218.
FR 17-NOV-2000; 2000US-0249224.
FR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249310.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0251160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251656.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.
XX
PA Rosen CA, Barash SC, Ruben SM;
P1 WPI; 2001-541565/60.
XX
DR
XX
ET Nucleic acids encoding 3224 human nervous system antigen polypeptides,
FT useful for preventing, diagnosing and/or treating nervous system
PT cancers and metastases -
XX
PS Disclosure; SEQ ID NO 12825; 1701pp + Sequence Listing; English.
XX

The invention relates to novel genes (ABAI11004-ABA2153) and proteins (ABBI4678-ABBI8001) useful for preventing, treating or ameliorating medical conditions e.g., by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital;

(b) immune disorders e.g. Addison's disease, allergies, autoimmune hemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Alignment Scores:		
Pred. No.:	5.55e-34	Length: 16877
Score:	359.00	Matches: 68
Percent Similarity:	100.00%	Conservative: 0
Best local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	22	Gaps: 0
US-09-830-144-4_COPY_437_504 (1-68) x ABA20494 (1-16877)		
Qy	1	GlnserProfrleuThrluEngInserThraanTrrHstInfrInserSerSerSer 20
Ddb	15011	CAAGGCCGACCTTAACCCGACGACTCCACCAACACGACGACGACGACGACTCCAGC 15070
Qy	21	SerSapglYglYleuPleatSerArpRoalAhIserleuProProglYgluaSapglY 40
Ddb	15071	TCTGACGAGGCTCTTCCGCTCCGCGCCGCGCCACTCGCTCCGCTCGCGGAGACGGT 15130

QY	41	ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTyrSerValAspHisGly	60
Db	15131	CGGTTGACCCCTATGATGGACTTTTGCTGAGATTTCACCCCTCTGAGCGTGGACCATGGC	15190
QY	61	GluGlnSerValValThrAlaPro	68
Db	15191	GAGCAGAGCGTGTGTACAGCACCG	15214
RESULT 9			
AL36984			
ID	AL36984	standard; DNA, 16877 BP.	
XX	AC		
XX	AA36984;		
DT	08-JAN-2002	(first entry)	
XX			
DE	Human musculoskeletal system related polynucleotide SEQ ID NO 3349.		
XX	KM	Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;	
XX	KM	antiallergic; hepatocytic; antidiabetic; antinflammatory; anticancer;	
XX	KM	valnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;	
XX	KM	cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder	
XX	KM	neurological disease; infection; human; secreted protein;	
XX	KM	musculoskeletal system; ds.	
OS	Homo sapiens.		
XX	PN	WO200155367-A1.	
XX	PD		
XX	02-AUG-2001.		
PF	17-JAN-2001;	2001WO-US01338.	
XX			
PR	31-JAN-2000;	2000US-0179065.	
PR	04-FEB-2000;	2000US-0180628.	
PR	24-FEB-2000;	2000US-0184664.	
PR	02-MAR-2000;	2000US-0186350.	
PR	16-MAR-2000;	2000US-0189874.	
PR	17-MAR-2000;	2000US-0190076.	
PR	18-APR-2000;	2000US-0198123.	
PR	19-MAY-2000;	2000US-0205515.	
PR	07-JUN-2000;	2000US-0209467.	
PR	28-JUN-2000;	2000US-0214886.	
PR	30-JUN-2000;	2000US-0215135.	
PR	07-JUL-2000;	2000US-0216647.	
PR	07-JUL-2000;	2000US-0216880.	
PR	11-JUL-2000;	2000US-0217487.	
PR	11-JUL-2000;	2000US-0217496.	
PR	14-JUL-2000;	2000US-0218290.	
PR	26-JUL-2000;	2000US-0220963.	
PR	26-JUL-2000;	2000US-0220964.	
PR	14-AUG-2000;	2000US-0224518.	
PR	14-AUG-2000;	2000US-0224519.	
PR	14-AUG-2000;	2000US-0225213.	
PR	14-AUG-2000;	2000US-0225214.	
PR	14-AUG-2000;	2000US-0225266.	
PR	14-AUG-2000;	2000US-0225267.	
PR	14-AUG-2000;	2000US-0225268.	
PR	14-AUG-2000;	2000US-0225270.	
PR	14-AUG-2000;	2000US-0225447.	
PR	14-AUG-2000;	2000US-0225757.	
PR	14-AUG-2000;	2000US-0225758.	
PR	14-AUG-2000;	2000US-0225759.	
PR	18-AUG-2000;	2000US-0226279.	
PR	22-AUG-2000;	2000US-0226681.	
PR	22-AUG-2000;	2000US-0226688.	
PR	23-AUG-2000;	2000US-0227182.	
PR	23-AUG-2000;	2000US-0227009.	
PR	30-AUG-2000;	2000US-0228994.	
PR	01-SEP-2000;	2000US-0229287.	
PR	01-SEP-2000;	2000US-0229343.	
PR	01-SEP-2000;	2000US-0229344.	
PR	01-SEP-2000;	2000US-0229345.	

Qy 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGly 40
Db 15071 TCTGACGAGGCGCTCTCCGCTCCGCGCCGACCTCCGCTCCGCTGCGAGGACGCT 15130
Qy 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60
Db 15131 CGTGTGAGCCCTATGAGACTTGTGAGATTACCGCTCTGAGAGCTGAGCATGAGC 15190
Qy 61 GluGlnSerValValThrAlaPro 68
Db 15191 GAGCAGAGCGGTGACAGCAGCCG 15214
RESULT 10
AAFI5895
ID AAFI5895 standard; cDNA; 696 BP.
XX
AC AAFI5895;
XX
DT 13-MAR-2001 (first entry)
XX
DE Human prostate cancer antigen nucleotide sequence SEQ ID NO:330.
XX
KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
KW neuroprotective; cytoskeletal; cardiovascular; immunomodulatory; muscular;
KW vulnerrary; gastrointestinal; nephrotoxic; antiinfective; gynaecological;
KW antibacterial; gene therapy; neural; immune; reproductive; renal;
KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
KW wound; infectious disease; ss.
XX
OS Homo sapiens.
XX
PN WO20005174-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US05988.
XX
PR 12-MAR-1999; 99US-0124270.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PA (ROSE/) ROSEN C A.
XX
PI Rosen CA, Ruben SM;
XX
WI WI: 2000-587513/55.
DR P-PSDB; AAB56692.
XX
PT Prostate cancer associated gene sequences, referred to as prostate
PT cancer antigens, useful for treatment, prevention, and diagnosis of
PT disorders such as prostate cancer -
XX
PS Claim 1; Page 837; 2338pp; English.
XX
CC AAFI566 to AAFI6505 encode the human prostate cancer associated
CC proteins, called prostate cancer antigens, given in AAB5633 to AAB57302.
CC The prostate cancer antigens can have neuroprotective, cytoskeletal,
CC cardiotropic, immunomodulatory, muscular, vulnerrary, gastrointestinal,
CC nephrotoxic, antiinfective, gynaecological and antibacterial activities,
CC and can be used in gene therapy. The prostate cancer antigen
CC polynucleotides may be used for detection of prostate cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The prostate cancer antigens may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal and proliferative
CC disorders, wounds, and infectious diseases. AAFI6506 to AAFI6514 to
CC AAB57303 represent sequences used in the exemplification of the present
CC invention.
XX
SQ Sequence 696 BP; 143 A; 237 C; 187 G; 120 T; 9 other;

Alignment Scores: 6 36e-35 Length: 696
Pred. No.: 352.00 Matches: 67
Score:

Percent Similarity: 98.53% Conservativity: 0
Best Local Similarity: 98.53% Mismatches: 1
Query Match: 98.05% Indels: 0
DB: 21 Gaps: 0
US-09-830-144-4_COPY_437_504 (1-68) x AAFI5895 (1-696)
Qy 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
Db 49 CAAAGCCGACCTTAACCTTGACAGTCCACACAGCAGCAGGAGGAGCAGCTCCAGC 108
Qy 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGly 40
Db 109 TCTRACGAGGCGCTCTCCGCTCCGCGCCGACCTGCTCCGCTGCGAGGACGCT 168
Qy 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60
Db 169 CGTGTGAGCCCTATGAGACTTGTGAGATTACCGCTCTGAGAGCTGAGCATGAGC 228
Qy 61 GluGlnSerValValThrAlaPro 68
Db 229 GAGCAGAGCGGTGACAGCAGCCG 252
RESULT 11
ID AABN61139/C
XX
AC AABN61139; standard; cDNA; 519 BP.
XX
DT 28-JUN-2002 (first entry)
XX
DE Human cancer related polynucleotide SEQ ID NO 1106.
XX
KW Human; cytoskeletal; gene expression; gene mapping; tissue profiling;
KW gene therapy; cancer; tumour; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200214500-A2.
XX
PD 21-FEB-2002.
XX
PF 16-AUG-2001; 2001WO-US25840.
XX
PR 16-AUG-2000; 2000US-226326P.
XX
PA (CHIR-) CHIRON CORP.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Escobedo J, Garcia PD, Sudduth-Klinger J, Reinhard C, Randazzo F;
PI Lamson G, Scott EM, Zhang G, Kassam A, Pot D, Labat I;
XX
WI WI: 2002-241905/29.
XX
PT New nucleic acid for producing a polypeptide, detecting differentially
PT expressed genes correlated with a cancerous state of a mammalian cell,
PT and inhibiting tumor growth -
XX
PS Claim 1; SEQ ID NO 1106; 883pp + Sequence Listing; English.
XX
CC The invention relates to an isolated polynucleotide (ABN27253-ABN33262)
CC with cytoskeletal activity. The polynucleotide is used to produce a
CC polypeptide, to detect differentially expressed genes correlated with a
CC cancerous state of a mammalian cell and to inhibit tumour growth. The
CC polynucleotide is used as a probe in mapping and tissue profiling. The
CC encoded polypeptide and antibodies to the polypeptide can also be used
CC for therapeutic and diagnostic purposes. The polynucleotide is useful for
CC gene therapy.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 519 BP; 152 A; 118 C; 114 G; 135 T; 0 other;

CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AA664197-AA694564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

SQ Sequence 990 BP; 215 A; 318 C; 287 G; 170 T; 0 other;

Alignment Scores:

Pred. No.:	8.18	Length:	990
Score:	74.50	Matches:	22
Percent Similarity:	38.96%	Conservative:	8
Best Local Similarity:	28.57%	Mismatches:	28
Query Match:	20.75%	Indels:	19
DB:	23	Gaps:	2

US-09-830-144-4_COPY_437_504 (1-68) x AA80813 (1-990)

QY 2 SerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSer 21
:::|||||:::|||||:::|||||:::|||||:::|||||
Db 412 GCGCCACAGCCCGGATCTGCGACCCGTCACCCACAGTCCTTGGGAGCTCTCA 471
QY 22 AspGlyGlyLeuPheArgProAlaHisSerLeuProGlyGlnAspGly--- 40
|||||:::|||||:::|||||:::|||||:::|||||
Db 472 GATGGG-----AGGAACAGGTGCCCCCAGATGCTCGAAAGTAACCATGGCCGC 522
QY 41 -----ArgValGlnProTyr 45
|||||:::|||||:::|||||:::|||||:::|||||
Db 523 CCTCATGGCAGCTCAGCTGTTTGGGTATTATTTCGATCTGAGAGTGAAGAGAAC 582
QY 46 ValAspPheAlaGluPheTyrArgLeuTyrSerValAspHisGlyGln 62
:::|||||:::|||||:::|||||:::|||||:::|||||
Db 583 ATTCCAGATGTGAAGACTTCAGAACGTGGACCTCGTTCGTGAGAGAG 633

RESULT 14

AAQ49396/C
ID AAQ49396 standard; DNA; 1108 BP.

XX	AAQ49396;
AC	27-APR-1994 (first entry)
DT	Class II AP endonuclease.
DE	Apurinic/apyrimidinic; recombinant; heat stable; ss.
XX	Thermus thermophilus.
OS	
XX	Key
FT	Location/Qualifiers
FT	114..926
FT	/*tag= a
XX	MO9320191-A.
XX	14-OCT-1993.
PD	
XX	30-MAR-1993; 93WO-US03035.
PF	
XX	31-MAR-1992; 92US-0860861.
PR	16-APR-1992; 92US-0869306.
XX	

PA (ABBO) ABBOTT LAB.
PA (CHAD) CHAKAKI DYEING KK.

XX Backman KC, Mclean PA;

XX WPI; 1993-336899/42.

DR P-PSDB; AAR42817.

XX New recombinant DNA encoding heat stable AP endonuclease - and

PT derived vectors, host cells and proteins, useful for nucleic acid

PT amplification by ligase chain reaction

XX Claim 4; Fig 1A; 42pp; English.

XX The sequence is that encoding Thermus thermophilus class II AP

CC (apurinic/apyrimidinic) endonuclease which retains activity when

CC subjected to elevated temperatures for the time necessary to

CC effect denaturation of double-stranded nucleic acids.

XX

SQ Sequence 1108 BP; 170 A; 379 C; 393 G; 166 T; 0 other;

US-09-830-144-4_COPY_437_504 (1-68) x AAQ49396 (1-1108)

QY 18 SerSerSerSerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGly 37
:::|||||:::|||||:::|||||:::|||||:::|||||
Db 203 AGCGTGAGCGCCGAGGCGGCTGCTTCACACGCCGCCGACGCTTTTCCCGC 144
QY 38 GluAspGlyArgValGluPro 44
|||||:::|||||:::|||||:::|||||:::|||||
Db 143 --GATGGAAAGGTGAACCC 126

RESULT 15

ABK78928
ID ABK78928 standard; DNA; 996 BP.

XX	ABK78928;
AC	13-AUG-2002 (first entry)
DT	Bacillus clausii genomic sequence tag (GST) #1771.
DE	Differential gene expression; genomic sequenced tag; GST;
XX	altered culture condition; environmental stress;
KW	physiological provocation; ds.
XX	Bacillus clausii.
OS	
XX	WO200229113-A2.
PN	11-APR-2002.
PD	
XX	05-OCT-2001; 2001WO-US31437.
PF	
XX	06-OCT-2000; 2000US-0680598.
PR	27-MAR-2001; 2001US-279526P.
XX	(NOVO) NOVOZYMES BIOTECH INC.
PA	(NOVO) NOVOZYMES AS.
XX	Berka R, Clausen IG;
PI	WPI; 2002-416684/44.
DR	
XX	Monitoring differential expression of several genes in first Bacillus
PT	cell relative to expression of same genes in one or more second


```
; EARLIER APPLICATION NUMBER: US No. 608712209/070,060
; EARLIER FILING DATE: 1998-04-30
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 5372
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-357-746-1

Alignment Scores:
Pred. No.: 119          Length: 5372
Score: 67.50           Matches: 23
Percent Similarity: 47.30%  Conservative: 12
Best Local Similarity: 31.08%  Mismatches: 20
Query Match: 18.80%        Indels: 19
DB: 3                    Gaps: 3

US-09-830-144-4_COPY_437_504 (1-68) x US-09-357-746-1 (1-5372)
QY 2 SerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSer 21
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 843 TCTAGTACAGGCTCTCTGCCCCACAAATACAAATACATCTGAGAGCAACA 902
QY 22 AspGly-----GlyLeuPheArgSerArgProAlaHis--- 32
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 903 TCTGATTAAATATCTCTTACTATATCTGAGGCTCAGGCCCTGAGCATTAATCTCT 962
QY 33 -----SerLeuProGlyGluAspGlyArgValGluProTyrValAspPhe 48
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 963 GTACTCAGAGCTCCCTGCTGCCCTGTTGGAGCAGAGAGATGAC----- 1007
QY 49 AlaGluPheTyrArgLeuTyrPseValAspHisGlyGluGln 62
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1008 --CAGCAGCGGCGAGTTTACTATGTAGATCATGTTGAGAAA 1046

RESULT 15
US-09-152-060-23
; Sequence 23, Application US/09152060
; Patent No. 6448230
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: P2003P1 US
; CURRENT APPLICATION NUMBER: US/09/152,060
; EARLIER FILING DATE: 1998-09-11
; EARLIER APPLICATION NUMBER: PCT/US98/04858
; EARLIER FILING DATE: 1998-03-12
; EARLIER APPLICATION NUMBER: 60/040,762
; EARLIER FILING DATE: 1997-03-14
; EARLIER APPLICATION NUMBER: 60/040,710
; EARLIER FILING DATE: 1997-03-14
; EARLIER APPLICATION NUMBER: 60/050,934
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,100
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,357
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/048,189
; EARLIER FILING DATE: 1997-05-30
; EARLIER APPLICATION NUMBER: 60/057,765
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/068,368
; EARLIER FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
; LENGTH: 1101
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-152-060-23
```

```
Alignment Scores:
Pred. No.: 48.5          Length: 1101
Score: 63.50           Matches: 27
Percent Similarity: 41.43%  Conservative: 2
Best Local Similarity: 38.57%  Mismatches: 30
Query Match: 17.69%        Indels: 11
DB: 4                    Gaps: 3

US-09-830-144-4_COPY_437_504 (1-68) x US-09-152-060-23 (1-1101)
QY 7 LeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSerAsp----- 22
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 576 CTGACAGCGGCGGACCTGCGGGGTCAAGCAGACAGCTTGCGAGCCTCGGACCC 635
QY 23 -----GlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 636 CGACGCGCAGCCGCTTTCAGGCTCGGCCCGGGGACTCTCTGCCATCCCGAGAGACCCGC 695
QY 41 ArgValGluProTyr-----ValAspPheAlaGluPheTyrArgLeuTyrPse 56
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 696 TACGTGAGCTGTATGTGTCGTCGAGCAATGCAGAGTTCAGATGCTGCGGAGCGAACA 755
QY 57 ---ValAspHisGlyGluGlnSerValVal 65
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 756 GCCGTGCTATCGGGTGGTGTGAGGTGTG 785

Search completed: December 10, 2002, 07:27:27
Job time : 72 secs
```

```
Db 694 TCTAGTACAGGCTCTCTCCGCCGACAAATACAAATACATCTCTGAAGGACACA 753
Qy 22 AspGly-----GlyLeuPheArgSerArgProAlaHis--- 32
Db 754 TCTGGATTAAATTCCTCTTACTATATCTGGAGGCTCAGGCCCTAGGCCATTAAATCTCT 813
Qy 33 -----SerLeuProProGlyGluAspGlyArgValGluProTyrValAspPhe 48
Db 814 GTAACCTCAAGCTCCCTTGCACCTGGTGGGAGCAGAGATGGAC----- 858
Qy 49 AlaGluPheTyrArgLeuTrpSerValAspHisGlyGluGln 62
Db 859 ---CAGCAGCGGCGAGTTTACTATGTAGATCATCTTGAGAAA 897
RESULT 12
US-09-357-746-2
; Sequence 2, Application US/09357746
; Patent No. 6087122
; GENERAL INFORMATION:
; APPLICANT: ZENECA Limited
; TITLE OF INVENTION: HUMAN E3 UBIQUITIN PROTEIN LIGASE
; FILE REFERENCE: PHM.70312.N1
; CURRENT APPLICATION NUMBER: US/09/357,746
; EARLIER FILING DATE: 1999-07-21
; EARLIER APPLICATION NUMBER: US No. 6087122 60/073,839
; EARLIER FILING DATE: 1998-02-05
; EARLIER APPLICATION NUMBER: US No. 608712209/070,060
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 2559
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-357-746-2
Alignment Scores:
Pred. No.: 45.6 Length: 2559
Score: 67.50 Matches: 23
Percent Similarity: 47.30% Conservative: 12
Best Local Similarity: 31.08% Mismatches: 20
Query Match: 18.80% Indels: 19
DB: 3 Gaps: 3
US-09-830-144-4_COPY_437_504 (1-68) x US-09-357-746-2 (1-2559)
Qy 2 SerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 21
Db 694 TCTAGTACAGGCTCTCTCCGCCGACAAATACAAATACATCTCTGAAGGACACA 753
Qy 22 AspGly-----GlyLeuPheArgSerArgProAlaHis--- 32
Db 754 TCTGGATTAAATTCCTCTTACTATATCTGGAGGCTCAGGCCCTAGGCCATTAAATCTCT 813
Qy 33 -----SerLeuProProGlyGluAspGlyArgValGluProTyrValAspPhe 48
Db 814 GTAACCTCAAGCTCCCTTGCACCTGGTGGGAGCAGAGATGGAC----- 858
Qy 49 AlaGluPheTyrArgLeuTrpSerValAspHisGlyGluGln 62
Db 859 ---CAGCAGCGGCGAGTTTACTATGTAGATCATCTTGAGAAA 897
RESULT 13
US-09-070-060-1
; Sequence 1, Application US/09070060
; Patent No. 5976849
; GENERAL INFORMATION:
; APPLICANT: Hustad, Carolyn M.
; APPLICANT: Chidyal, Namit
; TITLE OF INVENTION: Human E3 Ubiquitin Protein
; FILE REFERENCE: PHM.70312.N1
; CURRENT APPLICATION NUMBER: US/09/357,746
; EARLIER FILING DATE: 1999-07-21
; EARLIER APPLICATION NUMBER: US No. 6087122 60/073,839
; CORRESPONDENCE ADDRESS:
```

```
; ADDRESSEE: ZENECA Pharmaceuticals, Inc.
; STREET: 1800 Concord Pike
; CITY: Wilmington
; STATE: DE
; COUNTRY: USA
; ZIP: 19850-5437
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/070,060
; FILING DATE: 30-APR-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/073,839
; FILING DATE: 05-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Higgins, Patrick H
; REGISTRATION NUMBER: 39,709
; REFERENCE/DOCKET NUMBER: PHM.70312
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302.886.4889
; TELEFAX: 302.886.8221
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5359 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-070-060-1
Alignment Scores:
Pred. No.: 119 Length: 5359
Score: 67.50 Matches: 23
Percent Similarity: 47.30% Conservative: 12
Best Local Similarity: 31.08% Mismatches: 20
Query Match: 18.80% Indels: 19
DB: 3 Gaps: 3
US-09-830-144-4_COPY_437_504 (1-68) x US-09-070-060-1 (1-5359)
Qy 2 SerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 21
Db 848 TCTAGTACAGGCTCTCTCCGCCGACAAATACAAATACATCTCTGAAGGACACA 907
Qy 22 AspGly-----GlyLeuPheArgSerArgProAlaHis--- 32
Db 908 TCTGGATTAAATTCCTCTTACTATATCTGGAGGCTCAGGCCCTAGGCCATTAAATCTCT 967
Qy 33 -----SerLeuProProGlyGluAspGlyArgValGluProTyrValAspPhe 48
Db 968 GTAACCTCAAGCTCCCTTGCACCTGGTGGGAGCAGAGATGGAC----- 1012
Qy 49 AlaGluPheTyrArgLeuTrpSerValAspHisGlyGluGln 62
Db 1013 ---CAGCAGCGGCGAGTTTACTATGTAGATCATCTTGAGAAA 1051
RESULT 14
US-09-357-746-1
; Sequence 1, Application US/09357746
; Patent No. 6087122
; GENERAL INFORMATION:
; APPLICANT: ZENECA Limited
; TITLE OF INVENTION: HUMAN E3 UBIQUITIN PROTEIN LIGASE
; FILE REFERENCE: PHM.70312.N1
; CURRENT APPLICATION NUMBER: US/09/357,746
; CURRENT FILING DATE: 1999-07-21
; EARLIER APPLICATION NUMBER: US No. 6087122 60/073,839
; EARLIER FILING DATE: 1998-02-05
```

Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0
US-09-830-144-4_COPY_437_504 (1-68) x US-09-529-279-10 (1-1569)
Qy 1 GlnSerProThleuthleuGlnSerThraThrhsthrGlnSerSerSerSer 20
Db 1315 CAAGGCCGACCTTAACCTTGACGTCACACACGACGACGACGACCTCCACG 1374
Qy 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProProGlyGlyAspGly 40
Db 1375 TCTGACGAGAGGCTCTTCGCTCCGCGCCGCACTCGCTCCGCTGCGGAGGAGCGGT 1434
Qy 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTyrSerValAspHisGly 60
Db 1435 CGGTGTGAGCGCTATGTGACCTTGTCTGAGTTTACCGCTCTGAGACGTCATGCG 1494
Qy 61 GlnGlnSerValValThrAlaPro 68
Db 1495 GAGCAGAGCGTGTGACGACGACG 1518
RESULT 10
PCT-US93-03035-1/c
; Sequence 1, Application PC/TUS9303035
; GENERAL INFORMATION:
; APPLICANT: ABBOTT LABORATORIES
; TITLE OF INVENTION: PURIFIED THERMOSTABLE ENDONUCLEASE
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES
; STREET: D-377 ARD, ONE ABBOTT PARK ROAD
; CITY: ABBOTT PARK
; STATE: ILLINOIS
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/03035
; FILING DATE: 19930330
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/869,306
; FILING DATE: 16-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/860,702
; FILING DATE: 31-MAR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: BRAINARD, THOMAS D
; REGISTRATION NUMBER: 32,459
; REFERENCE/DOCKET NUMBER: 5145.PC.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-4884
; TELEFAX: 708-937-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1108 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 114..926
; PCT-US93-03035-1
Alignment Scores:
Pred. No.: 4.85 Length: 1108
Score: 71.50 Matches: 15

Percent Similarity: 70.37% Conservative: 4
Best Local Similarity: 55.56% Mismatches: 7
Query Match: 19.92% Indels: 1
DB: 5 Gaps: 1
US-09-830-144-4_COPY_437_504 (1-68) x PCT-US93-03035-1 (1-1108)
Qy 18 SerSerSerSerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProProGly 37
Db 203 AGCGTGAAGCGGAGGGGGTCTTCACCGCCCGCCACGCTTTTCCCGGC 144
Qy 38 GlnAspGlyArgValGluPro 44
Db 143 ---GATGAAAGGTGAACCC 126
RESULT 11
US-09-070-060-2
; Sequence 2, Application US/09070060
; Patent No. 5976849
; GENERAL INFORMATION:
; APPLICANT: Husted, Carolyn M.
; TITLE OF INVENTION: Human E3 Ubiquitin Protein
; TITLE OF INVENTION: Ligase
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZENECA Pharmaceuticals, Inc.
; STREET: 1800 Concord Pike
; CITY: Wilmington
; STATE: DE
; COUNTRY: USA
; ZIP: 19850-5437
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/070,060
; FILING DATE: 30-APR-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/073,839
; FILING DATE: 05-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Higgins, Patrick H
; REGISTRATION NUMBER: 39,709
; REFERENCE/DOCKET NUMBER: PHM.70312
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302.886.4889
; TELEFAX: 302.886.8221
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2559 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-070-060-2
Alignment Scores:
Pred. No.: 45.6 Length: 2559
Score: 67.50 Matches: 225
Percent Similarity: 47.30% Conservative: 12
Best Local Similarity: 31.08% Mismatches: 20
Query Match: 18.80% Indels: 19
DB: 2 Gaps: 3
US-09-830-144-4_COPY_437_504 (1-68) x US-09-070-060-2 (1-2559)
Qy 2 SerProThleuthleuGlnSerThraThrhsthrGlnSerSerSerSer 21
||| ||| ::||| |||||::||| ::|||

```
QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60
|||
Db 1458 CGTGTGAGCCCTATGTGAGCTTTGCTGAGTTTACCGCCCTCGGAGCGTGGACCATGCG 1517
|||
QY 61 GluGlnSerValValThrAlaPro 68
|||
Db 1518 GAGCAGAGCGTGTGACAGCACCG 1541
|||
RESULT 7
US-09-529-279-1
; Sequence 1, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT FILING DATE: 2000-04-11
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1560
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (30)..(1541)
US-09-529-279-1
Alignment Scores:
Pred. No.: 6.56e-36 Length: 1560
Score: 359.00 Matches: 68
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0
US-09-830-144-4_COPY_437_504 (1-68) x US-09-529-279-1 (1-1560)
QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
|||
Db 1338 CAAAGCCCGACCTTAACCTGCGTCCAGTCCACCAACGACGACGACGACGACGCTCCAGC 1397
|||
QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40
|||
Db 1398 TCTGACGAGGCGCTCTCCGCTCCCGCCGCCCACTCGCTCCGCTCGCGAGGACGCT 1457
|||
QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60
|||
Db 1458 CGTGTGAGCCCTATGTGAGCTTTGCTGAGTTTACCGCCCTCGGAGCGTGGACCATGCG 1517
|||
QY 61 GluGlnSerValValThrAlaPro 68
|||
Db 1518 GAGCAGAGCGTGTGACAGCACCG 1541
|||
RESULT 8
US-09-529-279-42
; Sequence 42, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT FILING DATE: 2000-04-11
```

```
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 42
; LENGTH: 1568
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11)..(1549)
US-09-529-279-42
Alignment Scores:
Pred. No.: 6.6e-36 Length: 1568
Score: 359.00 Matches: 68
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0
US-09-830-144-4_COPY_437_504 (1-68) x US-09-529-279-42 (1-1568)
QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
|||
Db 1346 CAAAGCCCGACCTTAACCTGCGTCCAGTCCACCAACGACGACGACGACGCTCCAGC 1405
|||
QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40
|||
Db 1406 TCTGACGAGGCGCTCTCCGCTCCCGCCGCCCACTCGCTCCGCTCGCGAGGACGCT 1465
|||
QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60
|||
Db 1466 CGTGTGAGCCCTATGTGAGCTTTGCTGAGTTTACCGCCCTCGGAGCGTGGACCATGCG 1525
|||
QY 61 GluGlnSerValValThrAlaPro 68
|||
Db 1526 GAGCAGAGCGTGTGACAGCACCG 1549
|||
RESULT 9
US-09-529-279-10
; Sequence 10, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT FILING DATE: 2000-04-11
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 1569
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)..(1557)
US-09-529-279-10
Alignment Scores:
Pred. No.: 6.61e-36 Length: 1569
Score: 359.00 Matches: 68
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
```


CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/752,891
FILING DATE: 20-NOV-1996
APPLICATION NUMBER: JP 8-300856
FILING DATE: 28-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 8-126282
FILING DATE: 24-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 17981/111
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1560 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 30..1541
NAME/KEY: mat_peptide
LOCATION: 30..1541
US-09-144-178-1

Alignment Scores:
Pred. No.: Length: 1560
Score: 359.00 Matches: 68
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-830-144-4_COPY_437_504 (1-68) x US-09-144-178-1 (1-1560)

QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
DB 1338 CAAGACCCGACCTTAACCTTCAGTCCACACACGACGACGAGCAGCTCCACG 1397
QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40
DB 1398 TCTGACGAGGCGCTTCTCCGCTCCGCGCCGCCACCTCGCTCCGCGAGGACGCT 1457
QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60
DB 1458 CGTGTGTGAGCCCTATGTGACATTTTGTGAGTTTACCGCCTCTGGAGCGTGGACCATGCG 1517
QY 61 GluGlnSerValThrAlaPro 68
DB 1518 GAGCAGAGCGTGGTGACAGCACCG 1541

RESULT 4
US-09-144-178-5
Sequence 5, Application US/09144178
Patent No. 5989862
GENERAL INFORMATION:
APPLICANT: MATSUMOTO, Kunihiro
APPLICANT: NISHIDA, Eisuke
TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/144,178
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/752,891
FILING DATE: 20-NOV-1996
APPLICATION NUMBER: JP 8-300856
FILING DATE: 28-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 8-126282
FILING DATE: 24-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 17981/111
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1560 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 30..1541
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 30..1541
US-09-144-178-5

Alignment Scores:
Pred. No.: Length: 1560
Score: 359.00 Matches: 68
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-830-144-4_COPY_437_504 (1-68) x US-09-144-178-5 (1-1560)

QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
DB 1338 CAAGACCCGACCTTAACCTTCAGTCCACACACGACGACGAGCAGCTCCACG 1397
QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40
DB 1398 TCTGACGAGGCGCTTCTCCGCTCCGCGCCGCCACCTCGCTCCGCGAGGACGCT 1457
QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60
DB 1458 CGTGTGTGAGCCCTATGTGACATTTTGTGAGTTTACCGCCTCTGGAGCGTGGACCATGCG 1517
QY 61 GluGlnSerValThrAlaPro 68
DB 1518 GAGCAGAGCGTGGTGACAGCACCG 1541

RESULT 5
US-09-406-854-1
Sequence 1, Application US/09406854
Patent No. 6140042
GENERAL INFORMATION:
APPLICANT: MATSUMOTO, Kunihiro
APPLICANT: NISHIDA, Eisuke
TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR

TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1560 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 30..1541
NAME/KEY: mat_peptide
LOCATION: 30..1541
US-08-752-891-1

Alignment Scores:
Pred. No.: 6,566-36 Length: 1560
Score: 359.00 Matches: 68
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-830-144-4_COPY_437_504 (1-68) x US-08-752-891-1 (1-1560)

Qy 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSer 20
Db 1338 CAAGGCCGACCTTAACCTTCAGTCCACCAACGACGACGACGACGACGACGACG 1397
Qy 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProProGlyGlyAspGly 40
Db 1398 TCTGACGAGGAGCTCTTCGCTCCGCGCCGACCTGCTCCGCTGCGAGAGAGG 1457
Qy 41 ArgValGlnProTyrValAspPheAlaGluPheTyrArgLeuTyrPserValAspHisGly 60
Db 1458 CGGTGTGAGCCCTATGATGAGCTTTGCTGAGTTTACCGCTCTGAGACGTCATGGC 1517
Qy 61 GlnGlnSerValValThrAlaPro 68
Db 1518 GAGCAGAGCGTGTGACAGCACCG 1541

RESULT 2
US-08-752-891-5
Sequence 5, Application US/08752891
Patent No. 5837819
GENERAL INFORMATION:
APPLICANT: MATSUMOTO, Kunihito
APPLICANT: NISHIDA, Eisuke
TITLE OF INVENTION: TABI PROTEIN AND DNA CODING THEREFOR
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/752,891
FILING DATE: 20-NOV-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 8-300856
FILING DATE: 28-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 8-126282

FILING DATE: 24-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 17981/111
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
TELEX: 904136

INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1560 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 30..1541
NAME/KEY: mat_peptide
LOCATION: 30..1541
US-08-752-891-5

Alignment Scores:
Pred. No.: 6,566-36 Length: 1560
Score: 359.00 Matches: 68
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-830-144-4_COPY_437_504 (1-68) x US-08-752-891-5 (1-1560)

Qy 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSer 20
Db 1338 CAAGGCCGACCTTAACCTTCAGTCCACCAACGACGACGACGACGACGACG 1397
Qy 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProProGlyGlyAspGly 40
Db 1398 TCTGACGAGGAGCTCTTCGCTCCGCGCCGACCTGCTCCGCTGCGAGAGAGG 1457
Qy 41 ArgValGlnProTyrValAspPheAlaGluPheTyrArgLeuTyrPserValAspHisGly 60
Db 1458 CGGTGTGAGCCCTATGATGAGCTTTGCTGAGTTTACCGCTCTGAGACGTCATGGC 1517
Qy 61 GlnGlnSerValValThrAlaPro 68
Db 1518 GAGCAGAGCGTGTGACAGCACCG 1541

RESULT 3
US-09-144-178-1
Sequence 1, Application US/09144178
Patent No. 5989862
GENERAL INFORMATION:
APPLICANT: MATSUMOTO, Kunihito
APPLICANT: NISHIDA, Eisuke
TITLE OF INVENTION: TABI PROTEIN AND DNA CODING THEREFOR
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/144,178
FILING DATE:

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 10, 2002, 05:36:06 ; Search time 68 Seconds

(without alignments)
306.677 Million cell updates/sec

Title: US-09-830-144-4_COPY_437_504

Perfect score: 359
Sequence: 1 QSPFLTQSTNTHQTSSSS.....AEFYRLWSVDHGRQSVVTAP 68

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool/US09830144/runat_04122002_141843_5006/app_query.fasta_1.263
-DB=Issued_Patents.NA -QWTF=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=BLOSUM62 -TRANS=numat40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09830144 @CGN_1_31 @runat_04122002_141843_5006 -NCPV=6 -ICPV=3
-NO_XLPXY -NO_MMAT -LARGQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents.NA.*

1: /cgn2_6/ptodata/1/ina/5A-COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B-COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A-COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B-COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS-COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	359	100.0	1560	2	US-08-752-891-1 Sequence 1, Appli
2	359	100.0	1560	2	US-08-752-891-5 Sequence 5, Appli
3	359	100.0	1560	2	US-09-144-178-1 Sequence 1, Appli
4	359	100.0	1560	2	US-09-144-178-5 Sequence 5, Appli
5	359	100.0	1560	3	US-09-406-854-1 Sequence 1, Appli
6	359	100.0	1560	3	US-09-406-854-5 Sequence 5, Appli
7	359	100.0	1560	4	US-09-529-279-1 Sequence 1, Appli
8	359	100.0	1568	4	US-09-529-279-42 Sequence 42, Appli
9	359	100.0	1569	4	US-09-529-279-10 Sequence 10, Appli
c 10	71.5	19.9	1108	5	PCT-US93-03035-1 Sequence 1, Appli
11	67.5	18.8	2559	2	US-09-070-060-2 Sequence 2, Appli
12	67.5	18.8	2559	3	US-09-357-746-2 Sequence 2, Appli

13	67.5	18.8	5359	2	US-09-070-060-1 Sequence 1, Appli
14	67.5	18.8	5372	3	US-09-357-746-1 Sequence 1, Appli
15	63.5	17.7	1101	4	US-09-152-060-23 Sequence 23, Appli
16	63.5	17.7	1173	3	US-08-706-218-5 Sequence 5, Appli
17	63.5	17.7	1473	4	US-09-152-060-43 Sequence 43, Appli
c 18	63.5	17.7	2784	4	US-08-857-076-39 Sequence 39, Appli
c 19	63.5	17.7	3017	4	US-08-857-076-52 Sequence 52, Appli
c 20	63.5	17.7	3119	4	US-08-857-076-53 Sequence 53, Appli
c 21	62	17.3	1395	2	US-08-553-367A-1 Sequence 1, Appli
c 22	62	17.3	1395	4	US-09-295-306-1 Sequence 1, Appli
c 23	62	17.3	1395	4	US-09-734-719-1 Sequence 1, Appli
c 24	62	17.3	5977	3	US-09-024-0208-1 Sequence 1, Appli
c 25	62	17.3	5977	4	US-09-425-043-1 Sequence 1, Appli
c 26	62	17.3	6007	3	US-09-024-0208-2 Sequence 2, Appli
c 27	62	17.3	6007	4	US-09-425-043-2 Sequence 2, Appli
c 28	62	17.3	6556	3	US-09-024-0208-7 Sequence 7, Appli
c 29	62	17.3	6556	4	US-09-425-043-7 Sequence 7, Appli
c 30	62	17.3	6586	3	US-09-024-0208-43 Sequence 43, Appli
c 31	62	17.3	6586	4	US-09-425-043-43 Sequence 43, Appli
c 32	62	17.3	6826	3	US-09-024-0208-8 Sequence 8, Appli
c 33	62	17.3	6826	4	US-09-425-043-8 Sequence 8, Appli
c 34	62	17.3	11303	4	US-08-961-527-115 Sequence 115, App
c 35	61.5	17.1	3478	1	US-08-530-492-1 Sequence 1, Appli
c 36	61.5	17.1	3478	4	US-08-906-517-1 Sequence 1, Appli
c 37	60.5	16.9	1518	2	US-08-997-080-88 Sequence 88, Appli
c 38	60.5	16.9	1518	2	US-08-997-362-88 Sequence 88, Appli
c 39	60.5	16.9	1518	3	US-08-873-970-88 Sequence 88, Appli
c 40	60.5	16.9	1518	4	US-09-095-855-88 Sequence 88, Appli
c 41	60.5	16.9	1518	4	US-09-324-542-88 Sequence 88, Appli
c 42	60.5	16.9	1518	4	US-09-205-426-88 Sequence 88, Appli
c 43	60.5	16.9	169998	4	US-09-676-6108-24 Sequence 24, Appli
c 44	60	16.7	903	5	PCT-US95-06406A-21 Sequence 21, Appli
c 45	60	16.7	1071	1	US-08-357-962-1 Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-08-752-891-1
; Sequence 1, Application US/08752891
; Patent No. 5837819
; GENERAL INFORMATION:
; APPLICANT: MATSUMOTO, Kunihiko
; APPLICANT: NISHIDA, Eisuke
; TITLE OF INVENTION: TABI PROTEIN AND DNA CODING THEREFOR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/752,891
; FILING DATE: 20-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-300856
; FILING DATE: 28-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-126282
; FILING DATE: 24-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 17981/111
; TELECOMMUNICATION INFORMATION:


```

Db 61 GGGGGCTTCCGCTCCAGACCGGCTCCTCCTCCTCCACCCGAGAGGATGCCGTGTG 120
Qy 43 GluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGlyGluGln 62
Db 121 GAGCCCTATGTGGACTTGTCTGAGTTCTTACCGACTCTGGAGCGTGGACCGAGGAGCAG 180
Qy 63 SerValValThrAlaPro 68
Db 181 AGCGTGATGACCGCACCT 198

RESULT 15
BI181203/c
LOCUS      BI181203              777 bp  mRNA      linear      EST 10-JUL-2001
DEFINITION UNL-P-FN-ae-a-05-0-UNL.s1 UNL-P-FN Sus scrofa cDNA clone
            UNL-P-FN-ae-a-05-0-UNL 3', mRNA sequence.
ACCESSION  BI181203
VERSION    BI181203.1 GI:14655612
KEYWORDS   EST.
SOURCE     pig.
ORGANISM   Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
            1 (bases 1 to 777)
Caetano, A.R., Johnson, R.K. and Pomp, D.
Generation and sequence characterization of a normalized cDNA
library from swine ovarian follicles
Unpublished (2001)
Contact: Pomp, D
Department of Animal Science
University of Nebraska, Lincoln
Lincoln, NE 68583-0908, USA
Tel: 402 472 6416
Fax: 402 472 6362
Email: dpomp@unl.edu
Oligo-dT track not found, Not 1 site shown in beginning of sequence
is likely internal to the message. The following repetitive
elements were found in this cDNA sequence: 61-111,
>GC rich#Low complexity
Seq primer: M13 -29
POLYA=NO.

FEATURES             Location/Qualifiers
     source           1..777
                     /organism="Sus scrofa"
                     /strain="University of Nebraska, Lincoln Swine Selection
                     Lines"
                     /db_xref="taxon:9823"
                     /clone_lib="UNL-P-FN"
                     /clone_lib="UNL-P-FN"
                     /dev_stage="ADULT"
                     /lab_host="DH10B (Life Technologies)"
                     /note="vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not 1; Site_2: Eco RI; The UNL-P-FN
                     library is a normalized library representing porcine
                     ovarian follicles, ranging between 2.0 to 10.0 mm in
                     diameter, collected during 7 days of the follicular phase
                     of the pig estrous cycle. This library was derived from
                     the library UNL-P-F2. The tag is a string of 5-6
                     nucleotides present between the Not 1 site and the
                     oligo-dT track. The library was constructed as described
                     by Bonaldo, Lennon and Soares, Genome Research 6: 791-806
                     , 1996.
BASE COUNT      106 a      247 c      290 g      127 t      7 others
ORIGIN

Alignment Scores:
Pred. No.:      3,82e-30      Length:      777
Score:          345.00      Matches:      65
Percent Similarity: 95.59%      Conservative: 0
Best Local Similarity: 95.59%      Mismatches: 3
Query Match:    96.10%      Indels:      0
DB:             13      Gaps:      0

```

US-09-830-144-4_COPY_437_504 (1-68) x BI181203 (1-777)

```

Qy 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSer 20
Db 317 CAGAGCCCGACCCCTGACCCCTGCAGTCCACCAACCCACCCAGCAGCAGCTCCAGC 258
Qy 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40
Db 257 TCCGACGGGGGCTCTTTTCGCTCCCGGCCACCCACTCGCTCCCGCCCGGCGGATGCG 198
Qy 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60
Db 197 CGCGTGGAGCCCTACGTGGACTTCGACAGAGTTCTACCGCTGTGGAGCGTGGACCGGC 138
Qy 61 GluGlnSerValValThrAlaPro 68
Db 137 GAGCAGAGTGTGGCGACGGGCGCG 114

```

Search completed: December 10, 2002, 07:26:07
Job time : 2220 secs

DEFINITION UNL-P-FN-au-g-01-0-UNL.s1 UNL-P-FN Sus scrofa cDNA clone
ACCESSION UNL-P-FN-au-g-01-0-UNL 3', mRNA sequence.
VERSION B1181306
KEYWORDS B1181306.1 GI:14655715
SOURCE EST.
ORGANISM pig.
REFERENCE Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
AUTHORS 1 (bases 1 to 815)
 Gaetano, A.R., Johnson, R.K. and Pomp, D.
TITLE Generation and sequence characterization of a normalized cDNA
 library from swine ovarian follicles
JOURNAL Unpublished (2001)
COMMENT Contact: Pomp, D
 Department of Animal Science
 University of Nebraska, Lincoln
 Lincoln, NE 68583-0908, USA
 Tel: 402 472 6416
 Fax: 402 472 6362
 Email: dpomp@unl.edu
 Oligo-dt track not found, Not I site shown in beginning of sequence
 is likely internal to the message. The following repetitive
 elements were found in this cDNA sequence: 62-112,
 >CC-rich#low_complexity
 Seq primer: M13 -29
 POLYA=NO.

FEATURES
source
 1..815
 /location/Qualifiers
 /organism="Sus scrofa"
 /strain="University of Nebraska, Lincoln Swine Selection
 Lines"
 /db_xref="taxon:9823"
 /clone="UNL-P-FN-au-g-01-0-UNL"
 /clone_lib="UNL-P-FN"
 /dev_stage="ADULT"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pMT73D-Pac (Pharmacia) with a modified
 polylinker, Site_1: Not I; Site_2: Eco RI; The UNL-P-FN
 library is a normalized library representing porcine
 ovarian follicles, ranging between 2.0 to 10.0 mm in
 diameter, collected during 7 days of the follicular phase
 of the pig estrous cycle. This library was derived from
 the library UNL-P-F2. The tag is a string of 5-6
 nucleotides present between the Not I site and the
 oligo-dt track. The library was constructed as described
 by Bonaldo, Lennon and Soares, Genome Research 6: 791-806
 , 1996.
 Tag_SEQ=None found"
 TAG_SEQ=266 c 306 g 134 t

BASE COUNT 109 a 266 c 306 g 134 t
ORIGIN

Alignment Scores:
 Pred. No.: 8.14e-31 Length: 815
 Score: 351.00 Matches: 66
 Percent Similarity: 97.06% Conservative: 0
 Best Local Similarity: 97.06% Mismatches: 2
 Query Match: 97.77% Indels: 0
 Gaps: 0

US-09-830-144-4_COPY_437_504 (1-68) x B1181306 (1-815)
 QY 1 GlnserProThleuThrluGlnserThraThnThiSthGlnserSerSerSerSer 20
 Db 318 CAGAGCCCGAGCCCTGAGCTCCAGTCCACCAACCCACACCCAGAGAGAGCTCCAGC 259
 QY 21 SerSpGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40
 Db 258 TCCACGAGGAGGAGCCCTCTTCCCTCCGCGCCACCCACTGCTCCGCGCGGAGAGAGTGC 199
 QY 41 ArgValGluProGlyValAlaPheAlaGluPheTyrArgLeuTyrSerValAlaHisGly 60
 Db 198 CGCGTGGAGGCGCTACGAGGAGCTTCGCGAGGTTCTACCGCGCTGTGGAGGTGAGACACGCGC 139

DEFINITION BM934363
LOCUS 644 bp mRNA linear EST 29-APR-2002
DEFINITION UI-M-CG0P-b11-a-03-0-UI.r1 NIH BMAP Ret4.S2 Mus musculus cDNA clone
 1 (bases 1 to 644)
ACCESSION UI-M-CG0P-b11-a-03-0-UI 5', mRNA sequence.
VERSION BM934363
KEYWORDS BM934363.1 GI:19393515
SOURCE EST.
ORGANISM house mouse.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS 1 (bases 1 to 644)
 Bonaldo, M.F., Lennon, G. and Soares, M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Chin, H
 National Institute of Mental Health
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
 20892-9643, USA
 Tel: 301 443 1706
 Fax: 301 443 9890
 Email: MEST@mail.nih.gov
 Tissue Procurement: Dr. Xin-Yuan Fu, Yale University School of
 Medicine
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seq primer: M13 REVERSE.

FEATURES
source
 1..644
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UI-M-CG0P-b11-a-03-0-UI"
 /clone_lib="NIH BMAP Ret4.S2"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pMT73D-Pac (Pharmacia) with a modified
 polylinker, Site_1: Not I; Site_2: Eco RI; The
 NIH_BMAP_Ret4.S2 library is a subtracted library,
 ultimately derived from mouse retina tissue libraries at
 various stages of development. For a detailed description
 of the library from which this clone was derived, please
 visit our web site at brainest.eng.uiowa.edu. The tissue
 for this library was contributed by Dr. Xin-Yuan Fu, Yale
 University School of Medicine"

BASE COUNT 138 a 210 c 174 g 121 t 1 others
ORIGIN

Alignment Scores:
 Pred. No.: 1.75e-30 Length: 644
 Score: 347.00 Matches: 65
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 98.48% Mismatches: 0
 Query Match: 96.66% Indels: 0
 Gaps: 0

US-09-830-144-4_COPY_437_504 (1-68) x BM934363 (1-644)
 QY 3 ProThleuThrluGlnserThraThnThiSthGlnserSerSerSerSer 22
 Db 1 CCACCTGACCTCGAGTCCACCAACGACGACACCCAGAGAGAGCTCCAGTGC 60
 QY 23 GlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGlyArgVal 42

Soares Lab, University of Iowa EST sequencing: M.B. Soares Lab, University of Iowa Clone distribution: clones will be available through Research Genetics (www.resgen.com) The following repetitive elements were found in this cDNA sequence: 62-112,
>GC rich#low complexity
Seq primer: M13 Forward
POLYA=No.

FEATURES

source

Location/Qualifiers
1..527
/organism="Sus scrofa"
/strain="crossbred"
/db_xref="taxon:9823"
/clone="MI-P-AV1-ncg-a-03-0-UI"
/clone_lib="MI-P-AV1"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not 1; Site_2: EcoRI; The MI-P-AV1 library is normalized library derived from the MI-P-AV0 library, ultimately derived from placenta tissue. For a detailed description of the library from which this clone was derived, please visit our web site at <http://bigest.genome.iastate.edu/>. The procedure used to create this library has been previously described (Bonaldo, Lennon and Soares, Genome Research 6: 791-806, 1996)
TAG_SEQ=None found"

BASE COUNT 67 a 166 c 216 g 78 t

ORIGIN

Alignment Scores:

Pred. No.: 4,63e-31 Length: 527
Score: 351.00 Matches: 66
Percent Similarity: 97.06% Conservative: 0
Best Local Similarity: 97.06% Mismatches: 2
Query Match: 97.77% Indels: 0
Gaps: 0

US-09-830-144-4_COPY_437_504 (1-68) x B1400511 (1-527)

QY 1 GlnSerProThleuthleuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
DB 318 CAGAGGCCCAACCTGACCTGACGATCCACCAACACCAACAGACAGACACCTCCACG 259
QY 21 SerASPGLVGLyleuPhearSerArpProAlaHisSerleuPProproGlyGluASPGLY 40
DB 258 TCCGAGCGGGGGCTCTTTCGCTCCGGGCCCAACCACTGCTCCCGCGCGAGATGGC 199
QY 41 ArgValGluProTyValAspPheAlaGluPheTyArgleuTpsSerValAspHisGly 60
DB 198 CGCGTGAGACCTACGTGACTTCGCGAGTTCTACCGCTGTGAGACGTGACACGCGC 139
QY 61 GluGlnSerValValThrAlaPro 68
DB 138 GAGCAGAGTGTGTGCGACGCGCGCG 115

RESULT 10 582 bp mRNA linear EST 17-SEP-2001
B1681846
LOCUS B1681846 582 bp mRNA linear EST 17-SEP-2001
DEFINITION 461334 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION B1681846
VERSION B1681846.1 GI:15634771
KEYWORDS EST.
SOURCE cow.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euteria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.

REFERENCE

AUTHORS

Smith,T.P.L., Grose,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
Cass,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,
G.U., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,
Perea,G., Holt,I., Karanymcheva,S., Liang,F., Quackenbush,J. and
Keefe,U.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA

JOURNAL libraries and construction of a gene index for cattle
MEDLINE Genome Res. 11 (4), 626-630 (2001)
COMMENT 21180013
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smithemil@marc.usda.gov
Single pass sequencing. Bases called and alt-trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCTGACGACG
Plate: 135 row: 0 column: 15
Seq primer: ATTAGGTGACACTATAG.

FEATURES

source

Location/Qualifiers
1..582
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="MARC 1BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPOR6; Site_1: NotI, Site_2: SalI,
Library made from pooled tissue from lymph node, ovary,
fat, hypothalamus, and pituitary."

BASE COUNT 93 a 212 c 179 g 98 t

ORIGIN

Alignment Scores:
Pred. No.: 5,26e-31 Length: 582
Score: 351.00 Matches: 67
Percent Similarity: 98.53% Conservative: 0
Best Local Similarity: 98.53% Mismatches: 1
Query Match: 97.77% Indels: 0
Gaps: 0

US-09-830-144-4_COPY_437_504 (1-68) x B1681846 (1-582)

QY 1 GlnSerProThleuthleuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
DB 80 CAGAGGCCCAACCTGACCTGACGATCCACCAACAGACAGACAGACACCTCCACG 139
QY 21 SerASPGLVGLyleuPhearSerArpProAlaHisSerleuPProproGlyGluASPGLY 40
DB 140 TCCGAGCGGGGGCTTTTCGCTCCGGGCCCAACCACTGCTCCCGCTGCGAGATGGC 199
QY 41 ArgValGluProTyValAspPheAlaGluPheTyArgleuTpsSerValAspHisGly 60
DB 200 CGCGTGAGACCTACGTGACTTCGCGAGTTCTACCGCTGTGAGACGTGACATGGC 259
QY 61 GluGlnSerValValThrAlaPro 68
DB 260 GAGCAGAGTGTGTGACGCGCGCG 283

RESULT 11 599 bp mRNA linear EST 14-AUG-2001
B1401530
LOCUS B1401530/C 599 bp mRNA linear EST 14-AUG-2001
DEFINITION MI-P-CP0-nvn-g-07-0-UI.s1 MI-P-CP0 Sus scrofa cDNA clone
ACCESSION MI-P-CP0-nvn-g-07-0-UI 3', mRNA sequence.
VERSION B1401530.1 GI:15180591
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euteria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE

AUTHORS

Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)

```

FEATURES
  source
    POLYA=No.
    Location/Qualifiers
      1..334
        /organism="Sus scrofa"
        /strain="crossbreed"
        /db_xref="taxon:9823"
        /clones="MI-P-Ay1-ngr-c-02-0-UI"
        /clone_lib="MI-P-Ay1"
        /lab_host="DH10B (Life Technologies)"
        /note="Vector: p773D-Pac (Pharmacia) with a modified
        polylinker; Site 1: Not I; Site 2: EcoRI; The MI-P-Ay1
        library is normalized library derived from the MI-P-Ay1
        library, ultimately derived from placenta tissue. For a
        detailed description of the library from which this clone
        was derived, please visit our web site at
        http://pigdb.genome.tastate.edu/. The procedure used to
        create this library has been previously described (Bonaldo
        , Lennon and Soares, Genome Research 6: 791-806, 1996)
        TAG_SEQ=None found"
      43 a _111 c 136 g 44 t
      BASE COUNT 43 a _111 c 136 g 44 t
      ORIGIN

Alignment Scores:
  Pred. No.: 2,56e-31 Length: 334
  Score: 351.00 Matches: 66
  Percent Similarity: 97.06% Conservative: 0
  Best Local Similarity: 97.06% Mismatches: 2
  Query Match: 97.77% Indels: 0
  DB: 12 Gaps: 0

US-09-830-144-4_COPY_437_504 (1-68) x BF710376 (1-334)
QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
Db 317 CAGAGCCCGACCTTCAGCTCCACCCACACCCACACCCAGAGCAGCTCCAGC 258
QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40
Db 257 TCCGACGGGGGCGCTCTTCGCTCCCGGCCACCCACCTCGCTCCCGCGCGGAGATGCG 198
QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60
Db 197 CGCGTGGAGCCCTAGTGGACTTCGCGGAGTTCTACCGCTCTGGAGCGTGGACACCGC 138
QY 61 GluGlnSerValValThrAlaPro 68
Db 137 GAGCAGAGTGTGGCGACGGCGCG 114

RESULT 8
LOCUS BG382061 378 bp mRNA linear EST 12-MAR-2001
DEFINITION 297803 MARC 1P1G Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BG382061.1 GI:13306533
VERSION BG382061.1
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE
AUTHORS Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.
and Keele,J.W.
TITLE Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine
JOURNAL Unpublished (2000)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@mail.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred

FEATURES
  source
    v0.980904.e. Vector identified by cross_match with the -minscore 18
    and -minmatch 12 options.
    PCR Primers
    FORWARD: AGGAAACAGCTATGACCAT
    BACKWARD: GTTTTCCAGTCAGCAGC
    Plate: 2 row: 0 column: 16
    Seq primer: ATTTAGTGACACTAG.
    Location/Qualifiers
      1..378
        /organism="Sus scrofa"
        /db_xref="taxon:9803"
        /clone_lib="MARC 1P1G"
        /tissue_type="pooled"
        /lab_host="DH10B"
        /note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
        Library made from pooled tissue from day 11, 13, 15, 20,
        and 30 embryos."
      52 a 154 c 123 g 49 t
      BASE COUNT 52 a 154 c 123 g 49 t
      ORIGIN

Alignment Scores:
  Pred. No.: 3,01e-31 Length: 378
  Score: 351.00 Matches: 66
  Percent Similarity: 97.06% Conservative: 0
  Best Local Similarity: 97.06% Mismatches: 2
  Query Match: 97.77% Indels: 0
  DB: 12 Gaps: 0

US-09-830-144-4_COPY_437_504 (1-68) x BG382061 (1-378)
QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
Db 71 CAGAGCCCGACCTTCAGCTCCACCCACACCCACACCCAGAGCAGCTCCAGC 130
QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40
Db 131 TCCGACGGGGGCGCTCTTCGCTCCCGGCCACCCACCTCGCTCCCGCGCGGAGATGCG 190
QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGly 60
Db 191 CGCGTGGAGCCCTAGTGGACTTCGCGGAGTTCTACCGCTCTGGAGCGTGGACACCGC 250
QY 61 GluGlnSerValValThrAlaPro 68
Db 251 GAGCAGAGTGTGGCGACGGCGCG 274

RESULT 9
LOCUS BI400511/c 527 bp mRNA linear EST 14-AUG-2001
DEFINITION MI-P-Ay1-ngr-a-03-0-UI.s1 MI-P-Ay1 Sus scrofa cDNA clone
ACCESSION BI400511
VERSION BI400511.1 GI:15179572
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
COMMENT MEDLINE 97044477
Contact: Tuggle CK
Molecular Genetics Laboratory, Department of Animal Science
Iowa State University
201 Kildee Hall, Ames, IA 50011-3150, USA
Tel: 5152944252
Fax: 5152942401
Email: cktuggle@iastate.edu
Oligo-dT track not found. Not I site shown in beginning of sequence
is likely internal to the message. cDNA Library Preparation: M.B.

```


Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Cloned through the I.M.A.G.E. Consortium information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LLAM766 row: e column: 05
 High quality sequence stop: 583.
 Location/Qualifiers
 1. .705
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:3593452"
 /clone_lib="NCI CGAP Maml"
 /tissue_type="tumor, biopsy sample"
 /dev_stage="10 months, virgin"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"

BASE COUNT 159 a 231 c 194 g 121 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.77e-31 Length: 705
 Score: 356.00 Matches: 67
 Percent Similarity: 100.00% Conservatives: 1
 Best Local Similarity: 98.53% Mismatches: 0
 Query Match: 99.16% Indels: 0
 DB: 10 Gaps: 0

US-09-830-144-4_COPY_437_504 (1-68) x BE375985 (1-705)
 Qy 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
 Db 132 CAGAGCCCCACTGTGACCTTGCAGTCCACCAACAGCACGCCAGCAGCAGCTCCAGC 191
 Qy 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGly 40
 Db 192 TCTGACGGGGCCCTTCCGCTCCAGCCGGCTCACTCACTCCACCGGAGAGATGC 251
 Qy 41 ArgValGluProTyroValAspPheAlaGluPheTyroArgLeuTrpSerValAspHisGly 60
 Db 252 CGTGTGAGGCCCTATGTGGACTTGTGTGAGTTCTACCGACTCTGGAGCGTGGACCGC 311
 Qy 61 GluGlnSerValThrAlaPro 68
 Db 312 GAGCAGAGCGTGATGAGCGCACT 335

RESULT 4
 B1253992
 LOCUS 602975130F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5114622 5', mRNA sequence.
 DEFINITION B1253992
 ACCESSION B1253992
 VERSION B1253992.1 GI:14805965
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 782)
 NIH-MGC <http://imgc.ncbi.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: Incyte Genomics, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Cloned through the I.M.A.G.E. Consortium information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LLAM1278 row: k column: 07
 High quality sequence stop: 657.
 Location/Qualifiers
 1. .782
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5114622"
 /clone_lib="NIH MGC 12"
 /tissue_type="cervical carcinoma cell line"
 /lab_host="DH10B"
 /note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.4 kb. Library prepared by Life Technologies."

BASE COUNT 162 a 288 c 214 g 118 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2.64e-31 Length: 782
 Score: 355.00 Matches: 67
 Percent Similarity: 98.53% Conservatives: 0
 Best Local Similarity: 98.53% Mismatches: 1
 Query Match: 98.89% Indels: 0
 DB: 13 Gaps: 0

US-09-830-144-4_COPY_437_504 (1-68) x B1253992 (1-782)
 Qy 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
 Db 306 CAAAGCCCGACCTTAACCTCGAGTCCACCAACAGCACGCCAGCAGCAGCTCCAGC 365
 Qy 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGly 40
 Db 366 TCTGACGGAGGGCTCTTCGGCTCCCGGCCCTGCCACTCGCTCCCGCTGGCAGGACGT 425
 Qy 41 ArgValGluProTyroValAspPheAlaGluPheTyroArgLeuTrpSerValAspHisGly 60
 Db 426 CGTGTGAGCCCTATGTGACTTTGTGAGTTTACCGCTCTGGAGCGTGGACCATGSC 485
 Qy 61 GluGlnSerValThrAlaPro 68
 Db 486 GAGCAGAGCGTGGTGACACCG 509

RESULT 5
 BQ925425
 LOCUS AGENCOURT_8801162 NIH_MGC_129 Mus musculus cDNA clone IMAGE:6310974 5', mRNA sequence.
 DEFINITION BQ925425
 ACCESSION BQ925425
 VERSION BQ925425.1 GI:22340456
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 1385)
 NIH-MGC <http://imgc.ncbi.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Susan L. Sullivan, PhD.
 CDNA Library Preparation: ResGen, Invitrogen Corp
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Cloned through the I.M.A.G.E. Consortium information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

JOURNAL
COMMENT
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCM/DTP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Place: LHC820 row: k column: 23
High quality sequence stop: 652.

FEATURES
Location/Qualifiers
1. 974
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_id="N1H_MGC_9"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: ovary; Vector: pOT97; Site 1: XhoI; Site 2: EcoRI; CDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAC(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 209 a 311 c 278 g 176 t
ORIGIN

Alignment Scores:
Pred. No.: 12e-31 Length: 974
Score: 359.00 Matches: 68
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-830-144-4_COPY_437_504 (1-68) X BE898567 (1-974)

QY 1 GlnSerProThLeuThLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSer 20
Db 433 CAAAGCCCGACCTTAACCTTCGAGTCACCAACGACGACGACGACGACGCTCCAGC 492
QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProProGlyGlyAspGly 40
Db 493 TCTGACGAGGAGCCTCTCCGCTCCGAGCCGCGCACTGCTCCGCGCGGAGAGCGGT 552
QY 41 ArgValGluProGlyValAlaAspPheAlaGluPheTyrrArgLeuTrpSerValAspHisGly 60
Db 553 CGTGTGAGCCCTATGTGAGCTTGTGAGTTTACCGCTCTGTGAGCGTGACCATGSC 612
QY 61 GluGlnSerValValThrAlaPro 68
Db 613 GAGCAGAGCGTGTGACAGCAGCG 636

RESULT 2
LOCUS AM258218 496 bp mRNA linear EST 03-APR-2000
DEFINITION ug31h07.y1 NCI CGAP MamM Mus musculus cDNA clone IMAGE:2811037 5'-similar to TR:Q15750 Q15750 TAKI BINDING PROTEIN.; mRNA sequence.
ACCESSION AM258218
VERSION AM258218.1 GI:6631199
KEYWORDS EST.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS 1 (bases 1 to 496)
TITLE NIH-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
COMMENT Unpublished (1997)

COMMENT
Other ESTs: ug31h07.x1
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/bdip/image/image.html
Seq primer: -40RP from Gibco
High quality sequence stop: 439.

FEATURES
Location/Qualifiers
1. 496
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_id="IMAGE:2811037"
/clone_id="NCI CGAP MamM"
/tissue_type="tumor, gross tissue"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NciI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Rubin Humphreys, NIH"

BASE COUNT 113 a 158 c 140 g 85 t
ORIGIN

Alignment Scores:
Pred. No.: 1.12e-31 Length: 496
Score: 356.00 Matches: 67
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.53% Mismatches: 0
Query Match: 99.16% Indels: 0
DB: Gaps: 0

US-09-830-144-4_COPY_437_504 (1-68) X AM258218 (1-496)

QY 1 GlnSerProThLeuThLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSer 20
Db 17 CAAAGCCCGACCTTAACCTTCGAGTCACCAACGACGACGACGACGCTCCAGC 76
QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProProGlyGlyAspGly 40
Db 77 TCTGACGAGGAGCCTCTCCGCTCCGAGCCGCGCACTGCTCCACCCGAGAGGTGC 136
QY 41 ArgValGluProGlyValAlaAspPheAlaGluPheTyrrArgLeuTrpSerValAspHisGly 60
Db 137 CGTGTGAGCCCTATGTGAGCTTGTGAGTTTACCGACTGTGAGCGTGACCGGC 196
QY 61 GluGlnSerValValThrAlaPro 68
Db 197 GAGCAGAGCGTGTGACAGCAGCG 220

RESULT 3
LOCUS BE375985 705 bp mRNA linear EST 21-JUL-2000
DEFINITION 601299419P1 NCI CGAP MamM Mus musculus cDNA clone IMAGE:3593452 5'-mRNA sequence.
ACCESSION BE375985
VERSION BE375985.1 GI:9321350
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS 1 (bases 1 to 705)
TITLE NIH-MGC http://mgc.ncbi.nlm.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 10, 2002, 05:34:46 ; Search time 2213 Seconds
(without alignments)
497.647 Million cell updates/sec

Title: US-09-830-144-4_COPY_437_504
Perfect score: 359
Sequence: 1 QSPFTTLOSTNTHQSSSSS.....AEFYRLWSVDHGQSVWTAP 68

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xlp
-Q2/cgn2_1/USPTO.spool/US09830144/runat_04122002_141843_4986/app_query.fasta_1.263
-DB=EST -QPMI=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09830144 @CGN_1_1_1716 @runat_04122002_141843_4986 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	359	100.0	974	12	BE988567	BE988567 601681494
2	356	99.2	496	10	AW258218	AW258218 uc31h07.y
3	356	99.2	705	10	BE375985	BE375985 601229419
4	355	98.9	782	13	BI253992	BI253992 602975130
5	355	98.9	1385	14	BQ925425	BQ925425 AGENCOURT
6	353	98.3	997	12	BG115732	BG115732 602317018
C 7	351	97.8	334	12	BF710376	BF710376 MI-P-AV1-
	351	97.8	378	12	EG382061	EG382061 297803 MA
C 9	351	97.8	527	13	BI400511	BI400511 MI-P-AV1-
	351	97.8	582	13	BI681846	BI681846 461334 MA
C 11	351	97.8	599	13	BI401530	BI401530 MI-P-CFO-
	351	97.8	648	13	BI184403	BI184403 UNL-P-FN-
C 13	351	97.8	815	13	BI181306	BI181306 UNL-P-FN-
	347	96.7	644	14	BM934363	BM934363 UI-M-CGOP
C 15	345	96.1	777	13	BI181203	BI181203 UNL-P-FN-
	327	91.1	919	12	BF163305	BF163305 601771849
C 17	322	83.7	337	12	BF712308	BF712308 MI-P-A3-a
	311	86.6	325	10	AW658281	AW658281 94183 MAR
C 19	309	86.1	617	12	EG710962	EG710962 pglin.pk0
	287	79.9	1107	12	BE902307	BE902307 601676585
C 21	275	76.6	1576	11	AK009321	AK009321 Mus muscu
	267	74.4	450	9	AA674170	AA674170 vp97a05.r
C 23	252	70.2	584	12	BE901639	BE901639 601677788
	251.5	70.1	610	14	BQ387712	BQ387712 NISC.mn25
C 25	245.5	68.4	818	12	BE746542	BE746542 601580107
	171.5	47.8	617	13	BI981448	BI981448 fu53e22.x
C 27	170	47.4	417	10	BE167965	BE167965 CM2-HT051
	159.5	44.4	1016	17	CGS021V1	CGS021V1 602810906
C 29	135	37.6	867	14	BQ431917	BQ431917 AGENCOURT
	105	29.2	1102	17	CGS03QBR	CGS03QBR Tetraodon
C 31	99	27.6	267	13	BM287700	BM287700 528490 MA
	97	27.0	501	9	AL602977	AL602977 DKF7p0661
C 33	84.5	23.5	546	9	AA707854	AA707854 zh24g12.s
	83	23.1	481	9	AA198507	AA198507 mul9e08.r
C 35	78	21.7	300	9	AA403881	AA403881 vd80a02.r
	78	21.7	863	17	CNS04AQ7	CNS04AQ7 Tetraodon
C 37	78	21.7	904	17	CNS02909	CNS02909 Tetraodon
	78	21.7	925	17	CNS022SP	CNS022SP Tetraodon
C 39	77	21.4	435	9	AL818083	AL818083 AL818083
	77	21.4	457	12	BG262490	BG262490 WHE0936_F
C 41	77	21.4	494	12	BG262699	BG262699 WHE0940_C
	76	21.2	563	17	AQ060318	AQ060318 CIT-HSP-2
C 43	75	20.9	461	9	AA218321	AA218321 mv74g07.r
	75	20.9	633	14	BU009263	BU009263 QGH9P14.y

ALIGNMENTS

RESULT 1
LOCUS BE988567 974 bp mRNA linear EST 29-SEP-2000
DEFINITION 601681494F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3551886 5', mRNA sequence.
ACCESSION BE988567
VERSION BE988567.1 GI:10365176
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 974)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)


```
; APPLICANT: Swanson, Ronald V.
; APPLICANT: Feldman, Robert A.
; APPLICANT: Schleper, Christa
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
; FILE REFERENCE: DCORP.002A
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: US/10/027,806
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2 42432
; LENGTH: 42432
; TYPE: DNA
; ORGANISM: Cenarchaeum symbiosum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(10421)
; NAME/KEY: CDS
; LOCATION: (10625)...(11434)
; NAME/KEY: CDS
; LOCATION: (11478)...(13046)
; NAME/KEY: CDS
; LOCATION: (13046)...(14620)
; NAME/KEY: CDS
; LOCATION: (23558)...(24862)
; NAME/KEY: CDS
; LOCATION: (24913)...(25728)
; NAME/KEY: CDS
; LOCATION: (26504)...(26981)
; NAME/KEY: CDS
; LOCATION: (29655)...(30491)
; NAME/KEY: CDS
; LOCATION: (34559)...(36067)
; NAME/KEY: CDS
; LOCATION: (37002)...(37403)
; NAME/KEY: CDS
; LOCATION: (37404)...(38282)
; NAME/KEY: CDS
; LOCATION: (39454)...(40572)
; US-10-027-806-2
Alignment Scores:
Pred. No.: 1.86e+03 Length: 42432
Score: 64.50 Matches: 25
Percent Similarity: 47.30% Conservative: 10
Best Local Similarity: 33.78% Mismatches: 20
Query Match: 17.97% Indels: 19
DB: 9 Gaps: 6
US-09-830-144-4_COPY_437_504 (1-68) x US-10-027-806-2 (1-42432)
Qy 2 SerProthrLeu---ThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer--- 19
Db 34335 GCCCCCACCTTCGGGACTATACAGTTGGATCGGAGCTTTGCAAGCTCCGCTGCAGCATG 34394
Qy 20 -----SerAspGlyGlyLeuPheArgSerArgPro---AlaHisSerLeuProPro 36
Db 34395 AAATATTATCCAGGTAGGG-----CGCTAAAGCGGCTCAGCGGCAGCGCTCCCGC 34448
Qy 37 Gly-----GluAspGlyArgValGluProTyrValAspPheAlaGlu 50
Db 34449 CCCGCGATATTATGATATATCGGGGCGGCGCTCCACCGCACCC----- 34496
Qy 51 PheTyrArgLeuTrpSerValAspHisGlyGluGlnSerVal 64
Db 34497 -----CGTATATGATCTCGGATCAGGGGGGTAGAAACCATTA 34532
RESULT 15
US-09-852-797-23
; Sequence 23, Application US/09852797
; Patent No. US20020172994A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 28 Human Secreted Proteins
; FILE REFERENCE: PZ003P2
; CURRENT APPLICATION NUMBER: US/09/852,797
; CURRENT FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: 60/265,583
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 09/152,060
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: PCT/US98/04858
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/040,762
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/040,710
; PRIOR FILING DATE: 1997-03-14
; PRIOR APPLICATION NUMBER: 60/050,934
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,100
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,357
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/048,189
; PRIOR FILING DATE: 1997-05-30
; PRIOR APPLICATION NUMBER: 60/057,765
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: 60/048,970
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
; LENGTH: 1101
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-852-797-23
Alignment Scores:
Pred. No.: 25 Length: 1101
Score: 63.50 Matches: 27
Percent Similarity: 41.43% Conservative: 2
Best Local Similarity: 38.57% Mismatches: 30
Query Match: 17.69% Indels: 11
DB: 9 Gaps: 3
US-09-830-144-4_COPY_437_504 (1-68) x US-09-852-797-23 (1-1101)
Qy 7 LeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSerSerSerSerSerSerAsp----- 22
Db 576 CTCACAGCGCGCGGACCTTCGCGGGTCAGCGACGACGCTGGGCGAGCCTCTCTGGGAGCCC 635
Qy 23 -----GlyGlyLeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGly 40
Db 636 CGGACGGGAGCGCTTTCAGGCTCGGGCGGGGAGCTCTCTGCCATCCCGAGACCCGCG 695
Qy 41 ArgValGluProTyr-----ValAspPheAlaGluPheTyrArgLeuTrpSer----- 56
Db 696 TACGTGGAGCTGATGTGGTCGTGGACAATGCAGAGTTCAGATGCTGGGGAGCGAAGCA 755
Qy 57 ----ValAspHisGlyGluGlnSerValVal 65
Db 756 GCCGTGCGTCACTCGGGTGTGGAGGTGGTG 785
Search completed: December 10, 2002, 07:29:10
Job time : 94 secs
```



```

Db 52 CAGACTCCGGCGCTCAGCTCCACAGCTTCAGAGTAGGAGTGGCCGAGGGCTGGCCAG 111
Qy 15 -----GInserSerSerSerSeraspGlyGlyLeuPheArgSerArgPro 30
Db 112 GGGCGGCTGCTGCTCCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 171
Qy 31 AlaHisSerLeuProGlyGlyLeuAspGly 40
Db 172 CATCACCACTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 201

RESULT 11
US-09-764-864-1682/c
; Sequence 1682, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT23
; CURRENT APPLICATION NUMBER: US/09/764,864
; PRIORITY FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1682
; LENGTH: 8918
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-864-1682

Alignment Scores:
Pred. No.: 93.2 Length: 8918
Score: 68.00 Matches: 19
Percent Similarity: 47.06% Conservative: 5
Best Local Similarity: 37.25% Mismatches: 13
Query Match: 18.94% Indels: 14
DB: 2 Gaps: 2

US-09-830-144-4_COPY_437_504 (1-68) x US-09-764-864-1682 (1-8918)

Qy 23 GlyGlyLeuPheArgSer---ArgProAlaHis----- 32
Db 7922 GGAGGATCTTCAGGAACGTAAGGAGTGAAGTTGAACATCGGTGTGCTTT 7863

Qy 33 -----SerLeuProGlyGlyLeuAspGlyArgValGluProTyrValAspPhe 48
Db 7862 TAACGACGACGACGACTTCTCCGGAGAGAAATGGAGGGTGAATAATTGTGCGTTTG 7803
Qy 49 AlaGluPheTyrArgLeuTyrSerValAspHis 59
Db 7802 GCGGGTTTGGCTCTCTTCATTAAGTATGATCAT 7770

RESULT 12
US-09-764-864-1683/c
; Sequence 1683, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT23
; CURRENT APPLICATION NUMBER: US/09/764,864
; PRIORITY FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1683
; LENGTH: 8919
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-864-1683

Alignment Scores:
Pred. No.: 93.2 Length: 8919
Score: 68.00 Matches: 19
Percent Similarity: 47.06% Conservative: 5
Best Local Similarity: 37.25% Mismatches: 13
Query Match: 18.94% Indels: 14
DB: 2 Gaps: 2

```

```

Percent Similarity: 47.06% Conservative: 5
Best Local Similarity: 37.25% Mismatches: 13
Query Match: 18.94% Indels: 14
DB: 10 Gaps: 2

US-09-830-144-4_COPY_437_504 (1-68) x US-09-764-864-1683 (1-8919)

Qy 23 GlyGlyLeuPheArgSer---ArgProAlaHis----- 32
Db 7923 GGAGGATCTTCAGGAACGTAAGGAGTGAAGTTGAACATCGGTGTGCTTT 7864

Qy 33 -----SerLeuProGlyGlyLeuAspGlyArgValGluProTyrValAspPhe 48
Db 7863 TAACGACGACGACGACTTCTCCGGAGAGAAATGGAGGGTGAATAATTGTGCGTTTG 7804
Qy 49 AlaGluPheTyrArgLeuTyrSerValAspHis 59
Db 7803 GCGGGTTTGGCTCTCTTCATTAAGTATGATCAT 7771

RESULT 13
US-09-938-842A-2555/c
; Sequence 2555, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIORITY FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2555
; LENGTH: 966
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2555

Alignment Scores:
Pred. No.: 15.8 Length: 966
Score: 64.50 Matches: 19
Percent Similarity: 53.33% Conservative: 13
Best Local Similarity: 31.67% Mismatches: 19
Query Match: 17.97% Indels: 9
DB: 2 Gaps: 2

US-09-830-144-4_COPY_437_504 (1-68) x US-09-938-842A-2555 (1-966)

Qy 4 ThrlleuThrlleuGlnSerThraen-----ThrlleThrlleSerSerSerSer 21
Db 865 ACTCTCTCTTACTTCTCAAAATCATCTTGCACGCTGACAAATGCGCTTGCAAGTGTCT 806

Qy 22 AspGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGlyLeuAspPhe 39
Db 805 AATCCTCATCTGTTCCAAACCTCAATGCAATCTTACACCTTTGCAAGTGTCT 746

Qy 40 -----GlyArgValGluProTyrValAspPheAlaGluPheTyrArgLeu 54
Db 745 CAGAGTATCATATCAAAAGATTGCCACTTCAATAGACTTCAGATCGCATCTCTG 686

RESULT 14
US-10-027-806-2
; Sequence 2, Application US/10027806
; Patent No. US20020160476A1
; GENERAL INFORMATION:

```



```
|||||
Db 1375 TCTGACGAGGCGCTCTTCCGCTCCCGGCGCCGACCTCGCTCCCGCTGGCGAGGAGCGT 1434
QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTyrPserValAspHisGly 60
Db 1435 CGTGTGAGCCCTATGTGAGACTTGTGAGTTTACCGCCTCTGAGACGTCGACCATGCGC 1494
QY 61 GluGlnSerValValThrAlaPro 68
Db 1495 GAGCAGAGCGTGTGACAGCACCG 1518

RESULT 6
US-09-764-877-3349
; Sequence 3349, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P0005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 3349
; LENGTH: 16877
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-3349

Alignment Scores:
Pred. No.: 1,82e-35 Length: 16877
Score: 359.00 Matches: 68
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
DB: 10

US-09-830-144-4_COPY_437_504 (1-68) x US-09-764-877-3349 (1-16877)

QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
Db 15011 CAAAGCCCGACCTTAACCTGACGTCACCAACGACGACGACGACGACGACGCTCCAGC 15070
QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProPogIyGluAspGly 40
Db 15071 TCTGACGAGGCGCTCTTCCGCTCCGCGCCGACCACTGCTCCGCTGGCGAGGAGCGT 15130
QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTyrPserValAspHisGly 60
Db 15131 CGTGTGAGCCCTATGTGAGACTTGTGAGTTTACCGCCTCTGAGACGTCGACCATGCGC 15190
QY 61 GluGlnSerValValThrAlaPro 68
Db 15191 GAGCAGAGCGTGTGACAGCACCG 15214

RESULT 7
US-09-925-300-330
; Sequence 330, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
```

```
; SEQ ID NO 330
; LENGTH: 696
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (643)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (657)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (685)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-300-330

Alignment Scores:
Pred. No.: 2.55e-36 Length: 696
Score: 352.00 Matches: 67
Percent Similarity: 98.53% Conservative: 0
Best Local Similarity: 98.53% Mismatches: 1
Query Match: 98.05% Indels: 0
Gaps: 0
DB: 10

US-09-830-144-4_COPY_437_504 (1-68) x US-09-925-300-330 (1-696)

QY 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
Db 49 CAAAGCCCGACCTTAACCTGACGTCACCAACGACGACGACGACGACGCTCCAGC 108
QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProPogIyGluAspGly 40
Db 109 TCTRACGAGGCGCTCTTCCGCTCCCGCGCCGACCACTGCTCCGCTGGCGAGGAGCGT 168
QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTyrPserValAspHisGly 60
Db 169 CGTGTGAGCCCTATGTGAGACTTGTGAGTTTACCGCCTCTGAGACGTCGACCATGCGC 228
QY 61 GluGlnSerValValThrAlaPro 68
Db 229 GAGCAGAGCGTGTGACAGCACCG 252

RESULT 8
US-09-974-300-6219
; Sequence 6219, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods for Monitoring Multiple Gene
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6219
; LENGTH: 996
; TYPE: DNA
; ORGANISM: Bacillus clausii
US-09-974-300-6219

Alignment Scores:
Pred. No.: 2.44 Length: 996
Score: 71.00 Matches: 19
Percent Similarity: 50.00% Conservative: 13
Best Local Similarity: 29.69% Mismatches: 24
Query Match: 19.78% Indels: 8
Gaps: 3
DB: 10
```

```
;
;   REGISTRATION NUMBER: 29,768
;   REFERENCE/DOCKET NUMBER: 17981/111
;
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (202)672-5300
;   TELEFAX: (202)672-5399
;   TELEX: 904136
;
;   INFORMATION FOR SEQ ID NO: 5:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1560 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
;
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 30..1541
;
;   FEATURE:
;   NAME/KEY: mat_peptide
;   LOCATION: 30..1541
;
;   SEQUENCE DESCRIPTION: SEQ ID NO: 5:
;
US-10-123-427-5
```

```
Alignment Scores:
Pred. No.: 9,078-37 Length: 1560
Score: 359.00 Matches: 68
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0
```

US-09-830-144-4_COPY_437_504 (1-68) x US-10-123-427-5 (1-1560)

```
Qy 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
|
Db 1338 CAAAGCCCGACCTTAACCCCTGCAGTCCACCAACAGCAGCAGCAGCAGCTCCAGC 1397
|
Qy 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40
|
Db 1398 TCTGACGAGCGCTCTTCCGCTCCCGCCGCCCACTCGCTCCCGCTGGCGAGGACGGT 1457
|
Qy 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTyrPsrValAspHisGly 60
|
Db 1458 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCCTCTGGAGCGTGACCATGCG 1517
|
Qy 61 GluGlnSerValValThrAlaPro 68
|
Db 1518 GAGCAGAGCGTGTGACAGCACCG 1541
```

```
RESULT 4
US-10-158-895-42
; Sequence 42, Application US/10158895
; Patent No. US20020155624A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: TSUCHIYA, MASAYUKI
; APPLICANT: OHTOMO, TOSHIHIKO
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 42
; LENGTH: 1568
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
```

```
;
;   LOCATION: (11)..(1549)
;
US-10-158-895-42
Alignment Scores:
Pred. No.: 9,128-37 Length: 1568
Score: 359.00 Matches: 68
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
;
US-09-830-144-4_COPY_437_504 (1-68) x US-10-158-895-42 (1-1568)
;
Qy 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
|
Db 1346 CAAAGCCCGACCTTAACCCCTGCAGTCCACCAACAGCAGCAGCAGCAGCTCCAGC 1405
|
Qy 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40
|
Db 1406 TCTGACGAGCGCTCTTCCGCTCCCGCCGCCCACTCGCTCCCGCTGGCGAGGACGGT 1465
|
Qy 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTyrPsrValAspHisGly 60
|
Db 1466 CGTGTGAGCCCTATGTGGACTTTGCTGAGTTTACCGCCTCTGGAGCGTGACCATGCG 1525
|
Qy 61 GluGlnSerValValThrAlaPro 68
|
Db 1526 GAGCAGAGCGTGTGACAGCACCG 1549
```

RESULT 5

```
US-10-158-895-10
; Sequence 10, Application US/10158895
; Patent No. US20020155624A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 10
; LENGTH: 1569
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)..(1557)
;
US-10-158-895-10
```

```
Alignment Scores:
Pred. No.: 9,138-37 Length: 1569
Score: 359.00 Matches: 68
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
```

US-09-830-144-4_COPY_437_504 (1-68) x US-10-158-895-10 (1-1569)

```
Qy 1 GlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 20
|
Db 1315 CAAAGCCCGACCTTAACCCCTGCAGTCCACCAACAGCAGCAGCAGCAGCTCCAGC 1374
|
Qy 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGly 40
|
```

Alignment Scores:
Pred. No.: 9.07e-37 Length: 1560
Score: 359.00 Matches: 68
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-830-144-4_COPY_437_504 (1-68) x US-10-158-895-1 (1-1560)

QY 1 GlnSerProThleuThrluInSerThraSnrHsiThrgInSerSerSerSer 20
Db 1338 CAAAGCCGACCTTAACCTTGACATCCCAACACGACACGACGACGACCTCCAGC 1397
QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProProGlyGlyAspGly 40
Db 1398 TCTGACGAGGAGGCTCTCCGCTCCGCGCCGCGCCACCTCGCTCCGCGCTGAGAGGAGCGGT 1457

QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTyrSerValAspHisGly 60
Db 1458 CGTGTGAGCCCTTAAGTGTGACATTTCTGAGTTTACCGCTCTTGAGACCTGAGCATGAC 1517
QY 61 GluGlnSerValAlaThrAlaPro 68
Db 1518 GAGCAGAGCGGTGATGACAGCACCG 1541

RESULT 2
US-10-123-427-1
Sequence 1, Application US/10123427
Patent No. US20020119525A1
GENERAL INFORMATION:
APPLICANT: MATSUMOTO, Kunihiko
NISHIDA, Eisuke
TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/123,427
FILING DATE: 17-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/406,854
FILING DATE: <Unknown>
APPLICATION NUMBER: US/08/752,891
FILING DATE: 20-NOV-1996
APPLICATION NUMBER: JP 8-300856
FILING DATE: 28-OCT-1996
APPLICATION NUMBER: JP 8-126282
FILING DATE: 24-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 17981/111
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1560 base pairs
TYPE: nucleic acid

STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 30..1541
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 30..1541
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-123-427-1

Alignment Scores:
Pred. No.: 9.07e-37 Length: 1560
Score: 359.00 Matches: 68
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-09-830-144-4_COPY_437_504 (1-68) x US-10-123-427-1 (1-1560)

QY 1 GlnSerProThleuThrluInSerThraSnrHsiThrgInSerSerSerSer 20
Db 1338 CAAAGCCGACCTTAACCTTGACATCCCAACACGACACGACGACGACCTCCAGC 1397
QY 21 SerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProProGlyGlyAspGly 40
Db 1398 TCTGACGAGGAGGCTCTCCGCTCCGCGCCGCGCCACCTCGCTCCGCGCTGAGAGGAGCGGT 1457

QY 41 ArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTyrSerValAspHisGly 60
Db 1458 CGTGTGAGCCCTTAAGTGTGACATTTCTGAGTTTACCGCTCTTGAGACCTGAGCATGAC 1517
QY 61 GluGlnSerValAlaThrAlaPro 68
Db 1518 GAGCAGAGCGGTGATGACAGCACCG 1541

RESULT 3
US-10-123-427-5
Sequence 5, Application US/10123427
Patent No. US20020119525A1
GENERAL INFORMATION:
APPLICANT: MATSUMOTO, Kunihiko
NISHIDA, Eisuke
TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/123,427
FILING DATE: 17-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/406,854
FILING DATE: <Unknown>
APPLICATION NUMBER: US/08/752,891
FILING DATE: 20-NOV-1996
APPLICATION NUMBER: JP 8-300856
FILING DATE: 28-OCT-1996
APPLICATION NUMBER: JP 8-126282
FILING DATE: 24-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 10, 2002, 05:51:41 ; Search time 86 Seconds

(without alignments)
308.319 Million cell updates/sec

Title: US-09-830-144-4_COPY_437_504

Perfect score: 359

Sequence: 1 OSPTLTQSTWTHQTSSSSS.....AEFYRLMSVDHGQSVWTAP 68

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Zgapop 6.0 , Zgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 350425 seqs, 194966369 residues

Total number of hits satisfying chosen parameters: 700850

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp

-Q=/cgn2_1/USPTO_spool/US09830144/runat_04122002_141844_5032/app_query.fasta_1.263

-DB=PublishedApplications NA -QFASTAP -SUFFIX=p2n.rnpb -MINMATCH=0.1

-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62

-TRANS-human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORES=ptc -THR MAX=100

-THR_MIN=0 -ALIGN=15 -MOD=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0

-MAXLEN=200000000 -USER=US09830144 @CGN 1 1 36 @runat_04122002_141844_5032

-NCFU=6 -ICPU=3 -NO_XLPXY -NO_WMAP -LARGQUERY -NEG SCORES=0 -WAIT -LONGLOG

-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAOP=10 -XGAPEXT=0.5 -FCGAPOP=6

-FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7-

Database :

Published Applications NA:

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:

2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:

3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:

4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:

5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:

6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:

7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:

8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:

9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:

10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:

11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:

12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:

13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:

14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	359	100.0	1560	9	US-10-158-895-1
2	359	100.0	1560	12	US-10-123-427-1
3	359	100.0	1560	12	US-10-123-427-5
4	359	100.0	1568	9	US-10-158-895-42

5	359	100.0	1569	9	US-10-158-895-10
6	359	100.0	16877	10	US-09-764-877-3349
7	352	98.1	696	10	US-09-925-300-330
8	71	19.8	996	10	US-09-974-300-6219
C 9	69.5	19.4	1492	9	US-09-941-831-7
C 10	68	18.9	273	10	US-09-864-761-22569
C 11	68	18.9	8918	10	US-09-764-864-1682
C 12	68	18.9	8919	10	US-09-764-864-1683
C 13	64.5	18.0	4232	9	US-09-338-842A-2555
14	64.5	18.0	42432	9	US-10-027-806-2
15	63.5	17.7	1101	9	US-09-852-797-23
16	63.5	17.7	1101	10	US-09-853-161-23
17	63.5	17.7	1101	10	US-09-852-659A-23
18	63.5	17.7	1473	9	US-09-852-797-43
19	63.5	17.7	1473	10	US-09-853-161-43
20	63.5	17.7	1473	10	US-09-852-659A-43
C 21	63.5	17.7	2784	10	US-09-205-658-39
C 22	63.5	17.7	2784	10	US-09-844-353A-39
C 23	63.5	17.7	3017	10	US-09-205-658-52
C 24	63.5	17.7	3017	10	US-09-844-353A-52
C 25	63.5	17.7	3119	10	US-09-205-658-53
C 26	63.5	17.7	3119	10	US-09-844-353A-53
27	63.5	17.7	3236	10	US-09-954-456-43
C 28	63.5	17.7	42450	10	US-09-815-048-3
29	62.5	17.4	167343	10	US-09-962-436-281
30	62.5	17.4	167343	10	US-09-964-824A-273
C 31	62	17.3	447	10	US-09-864-761-10872
C 32	62	17.3	1161	10	US-09-371-307-77
C 33	62	17.3	2637	10	US-09-815-242-9859
C 34	61.5	17.1	245	10	US-09-960-352-12511
35	61	17.0	1149	10	US-09-764-847-330
36	61	17.0	1450	10	US-09-764-853-186
37	61	17.0	3320	10	US-09-862-658-1
38	61	17.0	4536	10	US-09-801-368-249
C 39	60.5	16.9	172	10	US-09-864-761-27512
C 40	60	16.7	429	10	US-09-960-352-14704
41	60	16.7	509	9	US-09-764-868-255
42	60	16.7	510	9	US-09-764-868-566
C 43	60	16.7	584	10	US-09-864-761-15289
44	60	16.7	900	10	US-09-982-809-21
45	60	16.7	1606	10	US-09-853-386-55

ALIGNMENTS

RESULT 1
US-10-158-895-1
; Sequence 1, Application US/10158895
; Patent No. US20020155624A1
; GENERAL INFORMATION:
; APPLICANT: ONO KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1560
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (30)...(1541)
US-10-158-895-1